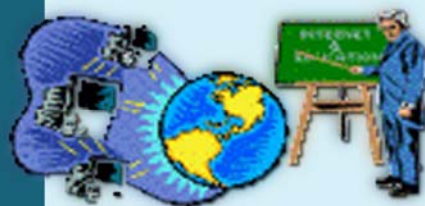


*Journal of*  
*Educational Technology*  
*& Society*

*Published by*  
*International Forum of Educational Technology & Society*



*October 2008*  
*Volume 11 Number 4*

# Educational Technology & Society

An International Journal

## Aims and Scope

*Educational Technology & Society* is a quarterly journal published in January, April, July and October. *Educational Technology & Society* seeks academic articles on the issues affecting the developers of educational systems and educators who implement and manage such systems. The articles should discuss the perspectives of both communities and their relation to each other:

- Educators aim to use technology to enhance individual learning as well as to achieve widespread education and expect the technology to blend with their individual approach to instruction. However, most educators are not fully aware of the benefits that may be obtained by proactively harnessing the available technologies and how they might be able to influence further developments through systematic feedback and suggestions.
- Educational system developers and artificial intelligence (AI) researchers are sometimes unaware of the needs and requirements of typical teachers, with a possible exception of those in the computer science domain. In transferring the notion of a 'user' from the human-computer interaction studies and assigning it to the 'student', the educator's role as the 'implementer/ manager/ user' of the technology has been forgotten.

The aim of the journal is to help them better understand each other's role in the overall process of education and how they may support each other. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to *Educational Technology & Society* and three months thereafter.

The scope of the journal is broad. Following list of topics is considered to be within the scope of the journal:

Architectures for Educational Technology Systems, Computer-Mediated Communication, Cooperative/ Collaborative Learning and Environments, Cultural Issues in Educational System development, Didactic/ Pedagogical Issues and Teaching/Learning Strategies, Distance Education/Learning, Distance Learning Systems, Distributed Learning Environments, Educational Multimedia, Evaluation, Human-Computer Interface (HCI) Issues, Hypermedia Systems/ Applications, Intelligent Learning/ Tutoring Environments, Interactive Learning Environments, Learning by Doing, Methodologies for Development of Educational Technology Systems, Multimedia Systems/ Applications, Network-Based Learning Environments, Online Education, Simulations for Learning, Web Based Instruction/ Training

## Editors

**Kinshuk**, Athabasca University, Canada; **Demetrios G Sampson**, University of Piraeus & ITI-CERTH, Greece; **Ashok Patel**, CAL Research & Software Engineering Centre, UK; **Reinhard Oppermann**, Fraunhofer Institut Angewandte Informationstechnik, Germany.

## Editorial Assistant

**Barbara Adamski**, Athabasca University, Canada.

## Associate editors

**Nian-Shing Chen**, National Sun Yat-sen University, Taiwan; **Alexandra I. Cristea**, Technical University Eindhoven, The Netherlands; **John Eklund**, Access Australia Co-operative Multimedia Centre, Australia; **Vladimir A Fomichov**, K. E. Tsiolkovsky Russian State Tech Univ, Russia; **Olga S Fomichova**, Studio "Culture, Ecology, and Foreign Languages", Russia; **Piet Kommers**, University of Twente, The Netherlands; **Chul-Hwan Lee**, Incheon National University of Education, Korea; **Brent Muirhead**, University of Phoenix Online, USA; **Erkki Sutinen**, University of Joensuu, Finland; **Vladimir Uskov**, Bradley University, USA.

## Advisory board

**Ignacio Aedo**, Universidad Carlos III de Madrid, Spain; **Luis Anido-Rifon**, University of Vigo, Spain; **Rosa Maria Bottino**, Consiglio Nazionale delle Ricerche, Italy; **Mark Bullen**, University of British Columbia, Canada; **Tak-Wai Chan**, National Central University, Taiwan; **Sherry Chen**, Brunel University, United Kingdom; **Darina Dicheva**, Winston-Salem State University, USA; **Robert Farrell**, IBM Research, USA; **Brian Garner**, Deakin University, Australia; **Roger Hartley**, Leeds University, UK; **J R Isaac**, National Institute of Information Technology, India; **Mohamed Jemni**, University of Tunis, Tunisia; **Paul Kirschner**, Open University of the Netherlands, The Netherlands; **William Klemm**, Texas A&M University, USA; **Rob Koper**, Open University of the Netherlands, The Netherlands; **Ruddy Lelouche**, Université Laval, Canada; **David McConnell**, Lancaster University, UK; **Rory McGreal**, Athabasca University, Canada; **David Merrill**, Brigham Young University - Hawaii, USA; **Marcelo Milrad**, Växjö University, Sweden; **Riichiro Mizoguchi**, Osaka University, Japan; **Permanand Mohan**, The University of the West Indies, Trinidad and Tobago; **Hiroaki Ogata**, Tokushima University, Japan; **Toshio Okamoto**, The University of Electro-Communications, Japan; **Thomas C. Reeves**, The University of Georgia, USA; **Gilly Salmon**, University of Leicester, United Kingdom; **Norbert M. Seel**, Albert-Ludwigs-University of Freiburg, Germany; **Timothy K. Shih**, Tamkang University, Taiwan; **Yoshiaki Shindo**, Nippon Institute of Technology, Japan; **Kevin Singley**, IBM Research, USA; **Brian K. Smith**, Pennsylvania State University, USA; **J. Michael Spector**, Florida State University, USA; **Chin-Chung Tsai**, National Taiwan University of Science and Technology, Taiwan; **Stephen J.H. Yang**, National Central University, Taiwan.

## Assistant Editors

**Sheng-Wen Hsieh**, Far East University, Taiwan; **Taiyu Lin**, Massey University, New Zealand; **Kathleen Luchini**, University of Michigan, USA; **Dorota Mularczyk**, Independent Researcher & Web Designer; **Carmen Padrón Nápoles**, Universidad Carlos III de Madrid, Spain; **Ali Fawaz Shareef**, Massey University, New Zealand; **Jarkko Suhonen**, University of Joensuu, Finland.

## Executive peer-reviewers

<http://www.ifets.info/>

## Subscription Prices and Ordering Information

For subscription information, please contact the editors at [kinshuk@iee.org](mailto:kinshuk@iee.org).

## Advertisements

*Educational Technology & Society* accepts advertisement of products and services of direct interest and usefulness to the readers of the journal, those involved in education and educational technology. Contact the editors at [kinshuk@iee.org](mailto:kinshuk@iee.org).

## Abstracting and Indexing

*Educational Technology & Society* is abstracted/indexed in Social Science Citation Index, Current Contents/Social & Behavioral Sciences, ISI Alerting Services, Social Scisearch, ACM Guide to Computing Literature, Australian DEST Register of Refereed Journals, Computing Reviews, DBLP, Educational Administration Abstracts, Educational Research Abstracts, Educational Technology Abstracts, Elsevier Bibliographic Databases, ERIC, Inspec, Technical Education & Training Abstracts, and VOCED.

## Guidelines for authors

Submissions are invited in the following categories:

- Peer reviewed publications: Full length articles (4000 - 7000 words)
- Book reviews
- Software reviews
- Website reviews

All peer review publications will be refereed in double-blind review process by at least two international reviewers with expertise in the relevant subject area. Book, Software and Website Reviews will not be reviewed, but the editors reserve the right to refuse or edit review.

For detailed information on how to format your submissions, please see:

<http://www.ifets.info/guide.php>

## Submission procedure

Authors, submitting articles for a particular special issue, should send their submissions directly to the appropriate Guest Editor. Guest Editors will advise the authors regarding submission procedure for the final version.

All submissions should be in electronic form. The editors will acknowledge the receipt of submission as soon as possible.

The preferred formats for submission are Word document and RTF, but editors will try their best for other formats too. For figures, GIF and JPEG (JPG) are the preferred formats. **Authors must supply separate figures** in one of these formats besides embedding in text.

Please provide following details with each submission: ■ Author(s) full name(s) including title(s), ■ Name of corresponding author, ■ Job title(s), ■ Organisation(s), ■ Full contact details of ALL authors including email address, postal address, telephone and fax numbers.

The submissions should be uploaded at [http://www.ifets.info/ets\\_journal/upload.php](http://www.ifets.info/ets_journal/upload.php). In case of difficulties, they can also be sent via email to (Subject: Submission for Educational Technology & Society journal): **kinshuk@ieee.org**. In the email, please state clearly that the manuscript is original material that has not been published, and is not being considered for publication elsewhere.

## Table of contents



### Full length articles

Visual Stereotypes and Virtual Pedagogical Agents <i>Magnus Haake and Agneta Gulz</i>	1-15
Communities of Ethical Practice: Using New Technologies for Ethical Dialectical Discourse <i>Linda Newman and John Findlay</i>	16-28
Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda <i>Tel Amiel and Thomas C. Reeves</i>	29-40
Randomised Items in Computer-based Tests: Russian Roulette in Assessment? <i>Anthony M. Marks and Johannes C. Cronje</i>	41-50
Investigation of Interaction, Online Support, Course Structure and Flexibility as the Contributing Factors to Students' Satisfaction in an Online Certificate Program <i>Erman Yukselturk and Zahide Yildirim</i>	51-65
A Two-Dimension Process in Explaining Learners' Collaborative Behaviors in CSCL <i>Jung-Lung Hsu, Huey-Wen Chou, Wu-Yuin Hwang and Shyan-Bin Chou</i>	66-80
Effects of Practice in a Linear and Non-linear Web-based Learning Environment <i>Florence Martin</i>	81-93
Content Recommendation Based on Education-Contextualized Browsing Events for Web-based Personalized Learning <i>Feng-Hsu Wang</i>	94-112
Using Teradata University Network (TUN), a Free Internet Resource for Teaching and Learning <i>Robert Winter, Anke Gericke and Tobias Bucher</i>	113-127
Hypergraphics for history teaching - Barriers for causal reasoning about history accounts <i>José M. Méndez and Manuel Montanero</i>	128-138
Developing an Intelligent Diagnosis and Assessment E-learning Tool for Introductory Programming <i>Chenn-Jung Huang, Chun-Hua Chen, Yun-Cheng Luo, Hong-Xin Chen and Yi-Ta Chuang</i>	139-157
Improving Recall and Transfer Skills Through Vocabulary Building in Web-Based Second Language Learning: An Examination by Item and Feedback Type <i>Seongchul Yun, Paul Chamness Miller, Youngkyun Baek, Jaeyeob Jung and Myunghwan Ko</i>	158-172
A Computer System of Referential Resolution to Assess Students' Reading Comprehension <i>Yu-Fen Yang, Wing-Kwong Wong and Hui-Chin Yeh</i>	173-189
Valuation of Online Continuing Medical Education and Telemedicine in Taiwan <i>Fuhmei Wang</i>	190-198
Supporting "Learning by Design" Activities Using Group Blogs <i>Georgios Fessakis, Konstantinos Tatsis and Angelique Dimitracopoulou</i>	199-212
A case study of enabling factors in the technology integration change process <i>Pi-Sui Hsu and Priya Sharma</i>	213-228
Teachers' Perceptions of the Dimensions and Implementation of Technology Leadership of Principals in Taiwanese Elementary Schools <i>I-Hua Chang, Joseph M. Chin and Cheng-Mei Hsu</i>	229-245
The Impact of Externalizing Readers' Mental Representation on the Comprehension of Online Texts <i>Yu-Fen Yang and Ya-Chen Hung</i>	246-264

A cross-cultural examination of the intention to use technology between Singaporean and Malaysian pre-service teachers: an application of the Technology Acceptance Model (TAM) <i>Timothy Teo, Wong Su Luan and Chai Ching Sing</i>	265-280
Educational Technology at a Crossroads: Examining the Development of the Academic Field in Canada <i>Rocci Luppigini</i>	281-296

## Visual Stereotypes and Virtual Pedagogical Agents

**Magnus Haake**

Dept. of Design Sciences, LTH, Lund University, P.O.Box 118, SE-221 00, Lund, Sweden //  
Tel: +46-(0)46-2227916 // Fax: +46-(0)46-2224431 // [magnus.haake@design.lth.se](mailto:magnus.haake@design.lth.se)

**Agneta Gulz**

LUCS (Div. of Cognitive Science), Lund University, Kungshuset, Lundagård, SE-222 22, Lund, Sweden //  
Tel: +46-(0)46-2223269 // Fax: +46-(0)46-2224424 // [agneta.gulz@lucs.lu.se](mailto:agneta.gulz@lucs.lu.se)

### ABSTRACT

The paper deals with the use of visual stereotypes in virtual pedagogical agents and its potential impact in digital learning environments. An analysis of the concept of visual stereotypes is followed by a discussion of affordances and drawbacks as to their use in the context of traditional media. Next, the paper explores whether virtual pedagogical characters introduce anything *novel* with regard to the use of visual stereotypes – as compared both to real life interaction between humans and to the use of visual stereotypes in traditional non-interactive media such as magazines, film, television and video. It is proposed that novel affordances, as well as novel drawbacks, indeed are being introduced with the use of visual stereotypes in virtual characters. The conclusion of the paper is that knowledge on these matters can be useful both for developers of educational systems and for educators in enabling them to strengthen some pedagogical settings and activities.

### Keywords

Pedagogical agent, Virtual character, Visual stereotype, Learning, Gender

### Introduction

Virtual pedagogical agents, i.e. computer generated characters in pedagogical roles are entering the digital society in increasing numbers. They are found in educational programs, from preschool to university. They are also found in broader educational contexts in the roles of virtual medical counsellors, physical exercise coaches and guides on city homepages, and they also appear in edutainment and infotainment settings. Furthermore, virtual characters are essential ingredients in digital environments used by young people, such as (of course) games and edutainment applications as well as chat systems and mobile phone applications (e.g. *imvu* at <http://www.imvu.com>).

According to Moreno & Flowerday (2006), we see more and more frequently that effective multimedia lessons are combined with the presence of virtual pedagogical agents. It is a qualified guess that various forms of digital educational environments and systems will increasingly incorporate these characters – virtual instructors, learning companions, coaches, mentors, etc. – concurrently with tool-kits for designing characters becoming increasingly available and wide-spread.

Furthermore, in recent years there has been a growing focus on the social dimension of human interaction with virtual pedagogical agents, with a variety of social competences being simulated in various agents (Baylor & Kim, 2005; Bickmore & Cassell, 2005; Gulz, 2005; Hall et al., 2004; Johnson, 2003; Paiva et al., 2004). Behind this lies the influential framework of “computers as social actors” by Reeves & Nass (1996) based on extensive studies that show that people spontaneously apply social interpretations and conventions when interacting with computer-based media. This human disposition, it is argued by Reeves & Nass and many others, ought to be further exploited in computer system design in order to make interaction smoother and more satisfying.

In human-human social interaction the *visual appearance* of other people is known to play a central role, with profound effects on our attitudes as well as behaviour. On the one hand, there are *dynamic visual aspects* such as gestures, facial expressions and gaze. These are extensively researched within the agent community. On the other hand, there are *static visual aspects* such as body and face properties, skin, hair and haircut, clothes and attributes. In spite of their documented impact in human-human social interaction (Kalick, 1988; McArthur, 1982), the static visual aspects – that we will focus on in this article – have been little attended to in research on virtual agents (Gulz & Haake, 2006a; Gulz & Haake, 2006b).

Particularly in interacting socially with unfamiliar others, humans exploit visual cues – static as well as dynamic – to form expectations for guiding the interaction. In drama theory the concept *physical personality* refers to the aspects of a drama character's appearance that immediately produce an impression of personality and initiate a set of expectations and attitudes. Among those are many static visual aspects such as body shape, height, sex, race, face, hair, clothing, make-up and facial hair (Brahnam, 2001). In contrast to how these cues are immediately picked up in the encounter with another human being, the linear stream of spoken information is incredibly slow (even though, of course, some voice characteristics are also quickly picked up). Berscheid & Walster (1974) note that, "... our appearance telegraphs more information about us than we would care to reveal on a battery of personality inventories, intelligence tests, and character scales. From flame-coloured hair through flat feet, few aspects of appearance fail to provide kernels of folk insight into another's nature." (ibid., p. 159).

These categorization processes, in which we quickly form expectations on a person's likely behaviour, attitudes, opinions, personality, manners, etc., rely heavily upon stereotypes – that is, ideas of "typical" representatives of certain human categories, where clusters of properties are ascribed to these categories and their representatives. Stereotype is one of the main concepts in this article, roughly standing for a culturally shared socio-cognitive schema that summarizes our beliefs of other people and acting as a default setting in social perception processes. Since our focus is on the visual static aspects of stereotypes, we will use the term *visual stereotypes*.

The objective of the article is to explore the use of visual stereotypes in virtual pedagogical agents and the potential impact of such use in educational technology with respect to teaching and learning. First, we ask to what extent affordances and drawbacks found in the use of visual stereotypes in real life interaction and in traditional media, reappear when used in virtual agents. In an attempt to answer this question, we also present and discuss some of the (few) empirical studies carried out in the area. Second, we analyse in what ways interaction with virtual pedagogical characters introduces something novel with respect to the use of visual stereotypes, both in terms of affordances and drawbacks. Before entering these two issues, we provide a background by analysing the concept of *visual stereotype*, relating it to the concept of *visual prototype*, and discussing visual stereotyping with respect to *visual naturalism* versus *visual stylisation*.

## The concept of visual stereotype

Figure 1 presents four examples of *visual stereotypes*. Many observers will in these pictures see a teenager (1a), a housewife (1b), a craftsman (1c), and an air-hostess (1d). A visual stereotype, in our sense of the term, consists of a number of visual attributes in a person that will make a majority of observers perceive the person as an illustration, or a typical instance, of a human group, a professional group, a social group, etc. That is, the visual input activates expectations on other – not visible – attributes in the person: how he/she is likely to behave and to talk, what he/she can be expected to say or not say, what attitudes and opinions he/she will be likely to have, etc. In this way visual cues carry social baggage.



Figure 1. Examples of visual stereotypes

*Gender* is often an important aspect of visual stereotypes. A “typical craftsman” is a man, whereas a “typical air hostess” is a woman (see Figure 1). Furthermore, a representation of a “typical scientist” is at the same time also a representation of a “typical male scientist”, whereas a “typical female scientist” is a stereotype of its own. Correspondingly, the visual stereotype of a “sweetie” concurs with that of a “female sweetie”, whereas “male sweetie” is another and separate visual stereotype. Gender will be a recurring issue in this text.

## Stereotypes and prototypes

Within the cognitive sciences the concept *prototype* is used to stand for “a typical exemplar” of a concept. For example, an apple is a prototypical fruit whereas a kiwi is not, and a lawyer is a prototypical law person whereas an investigation secretary is not. Focusing on the visual aspects, one may also speak of *visual prototypes* in parallel to *visual stereotypes*.

In the context of this article, an interesting difference between the concepts is that a (visual) prototype is something neutral, whereas a (visual) stereotype brings along negative associations (Schneider, 2003). A picture or description of a prototypical youngster, a prototypical nurse or a prototypical beauty is not necessarily something negative. The presumed basis for a certain exemplar being prototypical is the frequency of occurrences in human perceptions and experiences, and in that sense a prototype relates to something “really there”. A stereotype, on the other hand, is often associated with something that culture, media, etc. has constructed out of dubious starting points but where there is not something “really there”. (A complicating factor, that blurs the distinction just given, is of course that media constitutes an important part of our perceptions and experiences).

We will in the following use the term *visual stereotype* but attempt to assign to it some of the neutrality from the term prototype. Our purpose is, namely, to study the use of visual stereotypes from two angles and look for positive affordances as well as drawbacks. Thus, we are not assuming that the use of visual stereotypes is necessarily evil, unjustified or to be combated. (Actually, it was not until the 19<sup>th</sup> century that the word stereotype – originating from the Greek words *stereos* for “solid” and *typos* for “a model” – became linked to prejudice and discrimination (cf. Schneider, 2003).)

As to the bases of visual stereotypes, our standpoint is that it is a heterogeneous phenomenon. For some visual stereotypes the expression “no smoke without fire” applies. Culture and media may reinforce and exaggerate them, but there is “something there” behind the visual stereotypes. For instance, a punk look indeed often goes along with an individualistic attitude and a desire to be allowed to go one’s own way and not be forced to follow established societal norms. In the case of other visual stereotypes, the world has changed substantially compared to the situation from which they originate, and there is no or very little accuracy in them today: e.g. the personality and behavioural habits of the eccentric British explorer in his Khaki shorts, short-sleeved shirt, pith helmet, and brown shoes. Yet other visual stereotypes were from the start pure constructions. The “stupid blonde” stereotype is one such example. It certainly is not, nor has ever been, the case that blonde women (or men) on the average are stupider than non-blondes.

## Stylised visual stereotypes

Visual stereotypes can appear in many different media formats: photos, movies, paintings, drawings, comics, animated movies, etc. Some of these formats allow different degrees of visual naturalism, which in the context of stereotypes is a feature of interest. There is a whole scale from *photorealism* on the one hand to pronounced *stylisation* on the other, e.g. a cartoonish style or other artistic style that modifies and often simplifies a person’s appearance (Gulz & Haake, 2006b).

Stylisation makes it possible to sharpen and exaggerate a visual stereotype (see Figure 2). Such amplified visual stereotyping via stylisation is foremost associated with graphical media, such as cartoons and animated movies. But it is also used in theatre, where the shaping of characters sometimes makes extensive use of visual stereotypes. Heavy make-up and large distance to the stage can, furthermore, reinforce the appearance of the artists on the stage as visually stereotyped. In Commedia dell’Arte and classical Chinese opera, for example, dresses and make-up as well as gestures are pronouncedly stylised (see Figure 3). Also movies can rely upon stylised visual stereotyping. As



an example, Indiana Jones (Figure 3) wears throughout all three movies more or less the same outfit, characterized by his fedora, leather jacket, unbuttoned shirt and bullwhip, which function as immediate cues for identification. Likewise, it is often easy in movies to predict the roles of characters and who is going to die or survive on the basis of the visual stereotyping.



Figure 2. Visual stereotypes: naturalism versus stylization



Figure 3. Left: Harlequin (Commedia dell'Arte), middle: Chinese opera actor, and right: Indiana Jones (actor Harrison Ford)

## Visual stereotypes in traditional media characters as well as in virtual characters

### Visual stereotypes – affordances

Visual stereotypes are an important aspect of human thinking in their function as *cognitive short cuts* for making action and life tractable for human beings. Instead of becoming overwhelmingly occupied with thoughts and questions about people that we encounter we make use of their visual appearance to situate them, in order to focus on interaction as such (Brewer, 1988). That is, visual stereotypes frame our expectations. They are also used for building common references in conversations about other people. In brief, they are part of our social autopilot as an essential navigation tool in a social environment that would otherwise be overwhelmingly complex and demand a practically insurmountable burden of processing (Smith & Medin, 1981). Notably we speak here of very quick and largely unconscious processes.

If we start out to consider traditional visual media – theatre, film, comics – it is indeed essential that a reader or a spectator is scaffolded to gain an idea of the characters (their personalities, habits, manners, opinions,

predispositions, etc.). Without such starting points for entering the story, many plots would simply not be possible to follow. And here, as Laurel (1993) points out in discussing theatre, the visual appearance of characters can be used to suggest the internal traits of the character in order to function as shorthand for understanding and predicting the character. As mentioned above, some forms of theatre indeed drive this very far.

Consequently, a parallel use of visual stereotypes in character-based digital learning environments may provide starting points both for the interpretation of the virtual characters and for the interaction with them. Adequate starting points unleash resources to focus on the content of the interaction, afford smoother interaction and may generate a greater sense of enjoyment and accomplishment in users. This line of thought is indeed reflected in the many design recommendations or guidelines for virtual characters (as well as for traditional media characters), that underline the importance of *consistency* between features such as voice, gender, looks and role of a character (Nass et al., 2000). Inconsistencies engender disturbance and distraction, which increase the demand on cognitive resources. On the other hand, with a visual appearance that corresponds to behaviour and personality predictions, users get their expectations acknowledged and leveraged, and interaction becomes smooth and efficient.

However, choosing an *adequate visual stereotype* can be difficult as exemplified by the following three examples. In the first example, provided by de Rosis et al. (2004), a virtual character was to be designed for a digital legal information system in Italy. Initially the virtual character was modelled upon a very attractive young female assistant, since the developers assumed that the typical user of their system was going to be a male lawyer. However, after realizing that the lawyer's (female) secretary was the one who most frequently used the system, the designers became aware that the appearance and behaviour of the virtual character disturbed these users, and designed a new character with a more professional communication style and more classical attire. The point is that the first visual stereotype, the young attractive female secretary character, was not an adequate starting point for the users of the system – but instead disturbed and distracted them. Furthermore, *visual* redesign was indeed required. It would not have been sufficient to redesign the dialogue and behaviour of the character but leave the visual appearance intact.

A second example involves two different design cases of virtual assistants for city home pages. In Botkyrka, a Swedish community with a high percentage of immigrants, a stereotypically “Swedish-looking” light-blond female character was introduced as a virtual assistant on the website. However, due to negative comments from site visitors the character was removed and redesigned. This virtual character turned out *not* to be an adequate visual starting point for a visitor entering the site and its content. (In this case the content with its structure and pedagogical design may be perfectly well-designed but nevertheless of no benefit for visitors.) In contrast, in the Swedish city of Malmö, equally with a high percentage of immigrants, the character *Sara* (Figure 4) was chosen. In this character, stereotypical ethnicity attributes as well as stereotypical gender attributes, were treated carefully. Notably, the discussions underlying the visual design of the Sara character were extensive.



Figure 4. The virtual city guide Sara (<http://www.malmo.se/>)

A third example is provided by Baylor's research group at *Research of innovative technology for learning* (RITL). Within a programme for gender equity in science and engineering, Baylor directs the project *Challenging Stereotypes toward Engineering with Pedagogical Agents*. The project investigates, among other things, the influence of character appearance on female choice of engineering subjects. Several empirical studies within the project pointed in the same direction. Young female students were more strongly affected, as to their motivation and self-efficacy regarding engineering subjects, by virtual coaches that were similar to themselves or similar to how

they would like to be – in this case female, young and cool (Figure 5, left). Using such characters as coaches and instructors in tutorials on technology also seemed to increase the willingness of female students to select courses with technical content (Baylor, 2005). However, in an additional study (Baylor & Rosenberg-Kima, 2006) it was seen that even though, again, female students who interacted with a peer model character (female, young and cool) showed a more positive view of, and attitude towards, “an engineer”, the outcome was different as to influences on their attitudes regarding the importance and utility of engineering-related fields. Here the young female students were significantly more influenced by a male, older, un-cool “stereotypical engineer” character (Figure 5, right). The researchers tentative conclusion is that perhaps the “... most effective approach would be to use multiple agents (e.g., have a stereotypical engineer and a peer model both interact with participants).” (ibid., p. 6).



Figure 5. Virtual coaches from (Baylor & Rosenberg-Kima, 2006)

However, even if visual design choices can be difficult to make, we think it is important that educational systems developers acknowledge the influences of visual design, and in particular visual stereotypes. This may not come easily. Traditionally, computer science related domains have *not* acknowledged visual design as important in relation to the “real thing” which is the computer systems behind the – visually designed – interface. Of course the system must be wrapped up and presented in some way, but basically the system stands for itself. For virtual pedagogical agents, specifically, the argument goes that what really matters is the behaviour, the dialogue, the movements, the facial expressions and the pedagogical role of such an agent (Gulz & Haake, 2006a). The quality of these aspects determine to what extent the pedagogical goals set for the agent will be fulfilled. Then, indeed, the character needs to have a visual look or appearance – a skin colour a hair-cut, a body (shape), some clothing, etc. – but these are considered surface aspects with no real impact on the fulfilment pedagogical goals of the agent or of the learning environment that it inhabits.

This way of reasoning, we argue, is mistaken. On the one hand, it certainly is the case that without the development and refinement of algorithms and modules of behaviour, dialogue and pedagogy there would be *no virtual pedagogical agents at all*. On the other hand, no matter how adequate and well designed these fundamental aspects are – if the agent’s visual appearance is inadequate, the pedagogical benefits may decrease considerably. By now there is substantial evidence that learners’ expectations, attitudes, understanding and motivation in various ways are affected by the visual design of a virtual pedagogical agent. It may influence the following: learners’ beliefs in their own competence in approaching a certain subject matter, their willingness to pay attention to a presentation or tutorial, the extent to which they find something trustworthy or relevant, how hard they try to understand a material, and so on (Baylor & Plant, 2005; Graesser et al., 2004; Gulz et al., 2007b; Massaro, 2004; Moreno et al., 2001).

Furthermore, this does not refer only to “certain rare cases”, which involve explicit or apparent visual stereotypes. Stereotypes and stereotypical elements are pervasive in human cognition, and features such as gender, age, ethnicity, clothing, etc., must be decided on in almost all cases of a virtual character. After a series of studies involving such visual elements, Moreno & Flowerday (2006) argue that the choice of an agent’s visual appearance is practically always psychologically loaded, and put this empirically based conclusion in contrast to how “... the vast majority of instructional interfaces assign arbitrary animated pedagogical agents assuming that the choice of an agent representation is psychologically neutral.” (ibid., p. 191).

Even *subtle* cues as to visual stereotypes can influence learners’ experiences and the way they assimilate a given content in a digital context. This is shown, for instance, in (Gulz et al., 2007a) that presents a parallel study to one by Voelker (1994) and also described in Reeves & Nass (1996). The Voelker (1994) study compared user evaluations of

two female presenters, where one spoke in a more stereotypically *feminine* voice than the other. Results were that the presenter with the more feminine voice, and thus also the content of her presentation, was evaluated significantly lower on trustworthiness and intelligence, but significantly higher on warmth and empathy. In other words, relatively subtle voice cues evoked evaluations in line with well-known gender stereotypes. The present study (Gulz et al., 2007a) instead manipulated the degree of femininity of female virtual characters via *visual* cues. One character was designed as more (stereo)typically feminine and the other character as less (stereo)typically feminine by varying visual cues as to degree of femininity (shape of head, hairstyle and makeup) (Figure 6). The characters were also pre-validated in order to make sure that they were indeed perceived in this way. Regarding non-visual cues, both characters were identical as to their professional role as medical doctors, their voices and their lecturing on shift work and health. Again, the visual cues as to the degree of femininity influenced users' evaluation of the characters, and thus the content of their lectures, in accordance with gender stereotypes.



Figure 6. Virtual characters (presenters) from Gulz et al. (2007a)

It is important to remember that when *asked*, most people deny that one or another visual stereotype cue could make any difference for how they experience a presented material or for their attentiveness, etc. When told that this is the case, it can still be hard to believe or admit, possibly because it does not fit with the concept of human beings as rational and capable of identifying good content regardless of its form. Nevertheless, empirical evidence shows that we are all sensitive to visual stereotypes in the sense that on an *unconscious* cognitive level they influence our judgments and interpretations of informational settings that we encounter (Cook, 1979). Those involved in designing educational material and in teaching can profit from acknowledging and learning more about the mechanisms and processes involved.

### Visual stereotypes – drawbacks

A visual stereotype may – by its nature – activate misleading expectations. Even if based on some kind of frequency distributions of property-clusters in peoples' experiences and thus corresponding to “actuality” in a statistical sense (i.e. a prototype), a visual stereotype can in a given instance be inadequate and misleading. In real life an example could be a youngster whose street fashion look signals “tough, rebellious and cheeky” but who actually is very kind and helpful.

Correspondingly, a badly casted visual stereotype for a film character or a comic character – when not an *intentional* choice by the producers – may confuse and irritate users, and induce an impression of a non-believable and unprofessionally staged character. Consider all the fuss around the choice of James Bond actors, where some last for only one production, while others reappear again and again and become more or less synonymous to the role. Even though there are many different variables involved, the visual appearances play an important role in these outcomes.

For the case of interactive media, consider the Botkyrka example reported above, where the chosen visual stereotype of a “very Swedish-looking” women, activated undesired and misleading expectations as to the aim and use of the city home page (which was to welcome and invite everyone to use it, and to boost inhabitants' feelings of belonging to the city in question.)

Another important drawback is that visual stereotypes may be perceived to represent *the normal*, and make visual appearances that diverge from the stereotype be perceived as odd, unusual, or even abnormal. For instance, a



spectacled and somewhat thin craftsman is “no real craftsman”. In this way, visual stereotypes can hide or suppress nuances and an existing manifold. There is also the aspect of self-reproduction and self-reinforcing of stereotypes due to the close interactions between media and “real life. Societally undesirable gender stereotypes of a normative kind are frequent in traditional non-interactive media, and can be observed in the case of virtual characters as well. For instance, many computer game characters reproduce visual stereotypes. In 1998 it was concluded in the *Next Generation Magazine* that despite dramatic increases in the number of female game characters, “... they all seem to be constructed around very simple aesthetic stereotypes. In the East, it’s all giggling schoolgirls and sailor uniforms, but in the West the recipe appears to be bee-sting lips, a micro-thin waist, and voluminous, pneumatic breasts.” (Next Generation, 1998, p. 8). And even though there has been some change, overall there is still truth in this analysis. From more recent discussion forums one can learn that some female gamers refuse to play female characters and feel insulted by how they are designed, and also that some male gamers are unsatisfied with the masculine stereotypes presented (visually and otherwise): “I usually play as female characters, because male characters are always hyper-masculine and that’s not how I feel. However, if there’s a feminine guy, I WILL choose him.” (GameGirlAdvance, 2004).

In sum, we have discussed a number of advantages and drawbacks with the use of visual stereotypes, where what is known from traditional media reappears in the virtual world. We now proceed towards what virtual characters bring in as novel.

## Novel affordances and risks with virtual characters and visual stereotypes

### Introducing novel possibilities

All sets of visual cues that can appear in real human beings, or in photos and films of real humans, can also appear in a virtual agent. Thus, all visual stereotypes that can be seen in live human beings or in traditional media portraying humans can also be reproduced in virtual characters. But there are *additional possibilities* in virtual characters due to the extended degrees of freedom regarding visual modelling. In virtual characters it is easy to “cut and mix” and arrive at combinations that do not occur, or rarely so, in real human beings. Thus, it is relatively easy to challenge, or break down, visual stereotypes: to combine visual elements from different stereotypes or to combine a given visual stereotype with an unusual role.

It can be argued that such playing around with visual stereotypes is just as possible in other graphical media, such as comics and animated movies. Nevertheless it seems that in practice interactive media has brought this out more extensively. While traditional graphical media relies on the observer or reader as a passive consumer of pre-designed stereotypes, the interactive virtual arena activates the participants. In the area of computer gaming, we find communities where players themselves contribute to the design and development of characters. Here a remarkable character diversity can be observed. As to gender, several new appearances of female heroines, androgynous characters and other kinds of in-betweens have come into existence (Schleiner, 2000).

This points towards the potential of using virtual pedagogical characters as visual stereotype busters, to present the non-standardized and expose a manifold in combinations of ethnicity, professional roles, social classes, gender, and so on. Figure 7 shows some avatars from the on-line world *Second Life* (<http://secondlife.com>), of which some are used to explore alternative gender and personality.

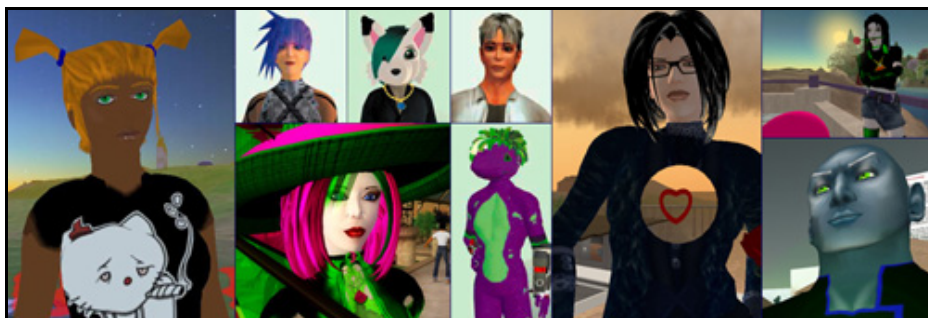


Figure 7. Avatars from *Second Life*

Offering a broader range of styles and identities may, furthermore, enable social identification and role modelling for a larger number of students. In turn, identification and opportunities for role modelling are known to strengthen development in pedagogical terms (Bandura, 1977). Also note the *specific* potentials of the digital virtual world. Here we can have dynamic, interactive situations, involving exploration and feedback that are not possible in traditional, non-interactive media. An example outside of education in a narrower sense, but belonging to the broader pedagogical domain, is a virtual character system for young women with eating disorders. The virtual coach character in question will start out looking really thin – in order for the clients to identify with, find trustworthy and be inclined to interact with – but then over time transform visually towards a more normal weight young female. In this sense, by slowly manipulating the stereotype, it may be possible to help the client to get away from an unhealthy mental visual stereotype.

But apart from the situation where *developers* may design virtual characters visually in a knowledgeable way in order to reach certain educational and pedagogical goals, one can also imagine the situation where learners themselves get to design their virtual instructors or learning companions. Where they, themselves or in a group, decide on the ethnicity, gender, body shape, clothing style, etc. for a virtual pedagogical character inhabiting a certain digital learning environment. “What is this instructor going to look like? Who is it going to be?” – Such a situation will be familiar for many young people used to games such as *the Sims*, avatars in on-line chats, etc. Furthermore, this situation can be set up both through commercial educational systems that provide character design kits, and through digital learning materials put together from scratch by teachers and students themselves, using character toolkits already becoming available (e.g. *PeoplePutty* at <http://www.haptek.com/peopleputty> and *Meez* at <http://www.meez.com>).

In both kinds of situations, a human pedagogue with knowledge of the impact of visual stereotypes will have an opportunity to use the situation as a basis for reflection and discussion: Why do we choose this character in this role – coach, learning companion, instructor, for different subject domains? Whose appearance shall be exhibited, in terms of gender, age, ethnicity, class, regional subgroup, etc? (Voice, in terms of gender, dialect and sociolect is another design feature that one will probably be able to choose). Which visual features do we find easy to combine and which not? What can be lost and what can be gained by choosing/designing a less naturalistic character?

Various alternatives may be suggested and explored. We believe that this kind of active and dynamic situation can provide a natural and powerful basis for reflection and discussion – more so than a standard and often more disconnected classroom discussion on stereotypes. This could be a rich and flexible tool for a pedagogue interested in challenging prejudices and proposing reconstruction of roles. Since humans are perceptual creatures and are powerfully affected by perceptual input and materials, this can be a vigorous complement to attempts at verbal reconstruction.

### **Dilemmas with novel possibilities**

The idea of *breaking with* visual stereotypes for pedagogical purposes that we have repeatedly lifted forth can, however, be in conflict with the pedagogical *exploiting of* visual stereotypes to facilitate smooth and efficient interaction. As observed earlier in the text, the use of visual stereotypes may enable learners to interact more smoothly with a character by acknowledging and leveraging learners’ expectations. In this way, the learner can focus on the learning activities and materials in question, rather than being confused and distracted by unexpected features and behaviour in a character.

The goal of smooth and efficient communication is central in the virtual agent research domain. In light of this, it is intelligible that Moreno et al. (2002) highlight the aim of obtaining “pedagogically effective animated agents” and the question of the “... role that stereotypic information [in the sense of visual stereotypes] plays in facilitating or inhibiting learning from animated agents.” (ibid., p. 4). Nevertheless, there is a striking lack of problematizing the issues and the results of their study, which indicate that participants learn significantly more from the male virtual tutors on the subject of blood pressure than from the female virtual tutors. The proposed explanation for this outcome goes that “... the female tutor broke with rules of etiquette about who should teach at a college level by not conforming to the stereotype of males as professors.” (ibid., p. 4), and is then left without further comment. That is, there is no mentioning of a conflict between, on the one hand, a wish to exploit the male professor stereotype in order to “facilitate learning from animated agents” and, on the other hand, a wish not to further reinforce the notion of the

male professor as the norm by using this stereotype. Likewise, the authors, without further comments, pose the question: "... do people learn more effectively about car repair from an agent named Joe who wears greasy overalls, or can they learn just as effectively about this topic from an agent named Nancy in a pink apron?" (ibid., p. 4). Compare Figure 8 showing an extract from the *Joe Doe* instruction series for US Army, drawn during the Second World War by Will Eisner.



Figure 8. Joe Dope (by Will Eisner)

In contrast to the Moreno et al. (2002) study, the complexity and hidden dilemmas in the use of visual stereotypes come forth clearly in the work of Baylor and her group at *RITL*. As related above, Baylor and collaborators have demonstrated (Baylor, 2005) that the use of virtual pedagogical coaches portrayed as young and attractive females can increase the willingness of female students to apply for technical education and to help increase their self-efficacy thanks to pedagogical processes such as role modelling and identification (cf. Bandura, 1977). The students find it easier to match these coaches compared to virtual coaches that are "typical, male, engineers", with their own personal identity. However a detailed analysis of the results indicate that the increase in self-efficacy at least partly stems from a conception of *such an engineer* – female, feminine, young and attractive – as less competent than a "real", prototypical, male engineer. What, according to the author, seems to occur is that the prejudice of females, and most of all feminine females, as less competent in technical domains spills over to the virtual area, generating increased self-efficacy of the kind "If she is able to do it, I can do it!". Now, this implies a potential conflict between a *short-term* pedagogical goal of recruitment and boosted self-efficacy in female students, and a *long-term* pedagogical goal of changing rather than reproducing gender prejudices and stereotypes. As Baylor (2005) remarks, prejudices about less competent female – especially feminine female – engineers are not ones that one would like to reinforce and disseminate.

Summing up, there is a need to handle dilemmas in which, on the one hand, the use of a visual stereotype can contribute to efficient communication in a pedagogical situation by leveraging users' expectations but where, on the other hand, the breaking with the stereotype can be desirable from a societal and long term pedagogical perspective. Likewise, there is a need to handle dilemmas where the breaking of a visual stereotype may have positive effects on attitudes as well as learning, but at the same time produce or reinforce questionable conceptions, such as the "female engineers are less capable" conception reported above.

### Introducing novel risks

As to the detrimental *normative* function of visual stereotypes discussed in a previous section, the extended degrees of design freedom offered in the virtual world brings about additional risks. The construction and promotion of *idealized super people* with "perfect" bodies and looks (and even lives) has long since been abounding in non-interactive media such as television, video and magazines. Figure 9 (right) shows the femme fatal *P'Gell* in Will

Eisner's comic *Spirit*. In this sense there is not much new under the sun, when interactive computer media continues this portraying of the ideal by promoting stereotypic instances never found in real life, such as the big-breasted, wasp-waisted action heroine *Lara Croft* (Figure 9, left): a biological contradiction whose tiny abdomen could hardly house all her vital organs, particularly if she also is to perform spectacular stunts and engage in violent fighting rather than fainting like corseted females of the 19<sup>th</sup> century.



Figure 9. Left: *Lara Croft* and right: *P'Gell* (by Will Eisner)

Nevertheless – this portrayal of the ideal can be taken *one step further* with interactive computer media. A key difference lies in what is otherwise seen as a central potential of virtual characters – not the least in pedagogical terms – namely their *interactivity*: Virtual characters may communicate, respond, and answer, thus establishing a dynamic, mutual social relation. A possible effect of this is that the distance between users or learners and these “ideal super people” is diminished. Until now we have watched, and read about, fabulous, good-looking people in movies and magazines (cf. P'Gell in Figure 9, right). If we are also to actively interact with them – in an era already desperately pursuing perfection in appearance – this might have detrimental effects on peoples self image and self esteem, as the interactivity may blur the distinction between “artefact” and “reality”.

Another related risk is the fact that virtual worlds involve users in a more active way than traditional movies and printed material. You can participate in various activities, including simulated everyday activities (c.f. *Second Life* at <http://secondlife.com> or *Entropia Universe* at <http://www.entropiauniverse.com>), and these activities may go on and on – there is no ending, as in the movie or book. This increases the risk (or potential) for users to indeed enter into “another world” with its characters and, in absorption, leave much of reality behind. The addiction risk is apparent. Extensive and absorbing interaction with stereotypical characters may have negative consequences for peoples' conceptions of real people and real social life.

## Conclusion

Positive and negative effects of visual stereotypes known from real life interaction and from traditional visual media reappear in interactive media. For instance, there is the invaluable function of visual stereotypes as cognitive tools for handling a complex social environment, as well as their problematic normative function that can make what diverges from a visual stereotype be perceived odd or abnormal.

But there are also *novel* possibilities and risks introduced by visual stereotypes in virtual agents. We have emphasized the extended possibilities to challenge visual stereotypes for educational purposes. Specifically we have highlighted the possibilities to provide social identification and role modelling for larger groups of learners. Rightly used, we think, virtual pedagogical agents can be a tool for supporting the exploration and formation of identity in



(young) learners while problemizing the reproduction of “undesired” (visual) stereotypes. Furthermore, it can be a pedagogical tool for initiating discussion and reflection on the role and the effects of visual stereotypes.

As to risks introduced by visual stereotypes in virtual agents, we have pointed at the replenished risks that follow from users *interacting* with, and perhaps being absorbed with, (visually) idealized stereotypes. The greater the knowledge about these issues among designers of digital learning material, the better the chance to counter these risks.

Summing up, we hold that the degrees of freedom as to visual design in digital virtual media, compared both to real life and to traditional media, imply an increased need for knowledge in order to navigate the design space in a thoughtful way. In products that are not directly commercial one may certainly wish that the visual designs of virtual pedagogical agents be based on informed design decisions.

### **Design guidelines**

This is a natural point to ask for *design guidelines* as a support for designers to produce adequate visual appearances for virtual pedagogical characters, and we indeed think that research results within the domain should be used to contribute to a “visual design guidelines project”. Such guidelines cannot, however, be step-by-step recipes on “how to visually design a virtual pedagogical agent”. The reason for this is that whether a visual design decision is adequate and appropriate will always depend also on the learning context, the learning goals and the group of learners in question. Yet guidelines in the form of pointers or topics and considerations to reflect upon in combination with good and bad examples are certainly both possible and desirable. Pointers suggesting that “this is a question that needs to be answered before doing a choice on this or that visual parameter” or that “these variables relate to each other” can support a designer’s navigation through the visual design space of virtual pedagogical agents.

Regarding visual stereotypes and expectations, there is also an overall design consideration and trade-off discussed earlier in the text that should be kept in mind. On the one hand, there are times when pedagogical benefits are gained by *challenging* pre-conceptions and pre-knowledge. On the other hand, there are times when there are pedagogical benefits in *exploiting* existing conceptions, expectations and preferences in students.

### **In sum**

We began by analysing the use of visual stereotypes in traditional, non-interactive, media, and indeed, there is interesting research in the case of traditional media relating to this topic. However, we hold that the issues must be separately approached for interactive pedagogical media, since there are additional affordances, in positive and negative senses, in the case of interactive media. By this standpoint we disagree to some extent with Reeves & Nass (1996) who hold that traditional media, such as television, and new media, such as computers “... afford the same problems and opportunities of stereotyping.” (ibid., p. 170). We believe that certain novel problems as well as opportunities enter the scene with digital interactive media. And if academic research in the domain of virtual pedagogical agents keeps up with the technological and commercial development, there is a potential to take on some responsibility and to be proactive in channelling the development.

### **Acknowledgements**

The authors are grateful for the financial support of the Swedish Research Council, VR. Also sincere thanks to Daniel Labbé and to anonymous reviewers for many valuable comments on a previous version of the article.

### **Copyrights**

**Figure 1:** Visual stereotypes © 2006 Magnus Haake.

**Figure 2:** Visual stereotypes © 2006 Magnus Haake.

**Figure 3 (left):** Harlequin © 2008 Magnus Haake. (Illustration based on a photograph by Baron Adolf de Meyer, portraying the Russian dancer Vaslav Nijinsky as Harlequin in the ballet *Carnaval*, Paris, 1910, Roger Pryer Dodge Collection, The New York Public Library.)

**Figure 3 (middle):** Chinese opera actor © 2008 Duc (www.pixiduc.tk).

**Figure 3 (right):** Actor Harrison Ford as Indiana Jones, COURTESY OF LUCASFILM LTD. *Indiana Jones and the Temple of Doom*™ © 1989 Lucasfilm Ltd. All rights reserved. Used under authorization. Unauthorized duplication is a violation of applicable law.

**Figure 4:** The webbot Sara. Information service at the homepage of City of Malmö, Sweden, www.malmo.se. Courtesy of City of Malmö, Sweden.

**Figure 5:** Virtual coaches from: Baylor, A. & Rosenberg-Kima, R. (2006). Interface agents to alleviate online frustration. Paper presented at the 7th International Conference of the Learning Sciences, Bloomington, IN. © 2004 RITL@FSU. All Rights Reserved.

**Figure 6:** Virtual characters (presenters) from: Gulz, A., Ahlner, F. & Haake, M. (2007a). Visual femininity and masculinity in synthetic characters & patterns of affect. In Proceedings of ACII 2007. Berlin/Heidelberg, Germany: Springer, 654-665. © 2006 Magnus Haake.

**Figure 7:** Screenshots from Second Life® © Linden Research, Inc.

**Figure 8:** Joe Doe instruction series for the US Army, Courtesy of Will Eisner Studios, Inc. Joe Dope™ © Will Eisner Studios, Inc. All rights reserved.

**Figure 9 (left):** Action heroine Lara Croft, Courtesy of Eidos Interactive Ltd. Lara Croft™ © Eidos Interactive Ltd. All rights reserved.

**Figure 9 (right):** Cartoon femme fatal P’Gell, Courtesy of Will Eisner Studios, Inc. P’Gell™ © Will Eisner Studios, Inc. All rights reserved.

## References

Bandura, A. (1977). *Social Learning Theory*, Englewood Cliffs, NJ: Prentice-Hall.

Baylor, A. (2005). The impact of pedagogical agent image on affective outcomes. *Paper presented at the 2005 International Conference on Intelligent User Interfaces (IUI 2005) Workshop on Affective Interactions: Computers in the Affective Loop*, January 9-12, 2005, San Diego, CA, retrieved June 11, 2007 from [http://garnet.acns.fsu.edu/~abaylor/PDF/image\\_iui.pdf](http://garnet.acns.fsu.edu/~abaylor/PDF/image_iui.pdf).

Baylor, A. & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, 15 (5), 95-115.

Baylor, A. & Plant, A. (2005). Pedagogical agents as social models for engineering: The influence of appearance on female choice. In Looi, C-K., McCalla, G., Bredeweg, B. & Breuker J. (Eds.), *Artificial intelligence in education: Supporting Learning through intelligent and socially informed technology*, 125, Amsterdam, The Netherlands: IOS Press, 65-72.

Baylor, A. & Rosenberg-Kima, R. (2006). Interface agents to alleviate online frustration. *Paper presented at the 7th International Conference of the Learning Sciences*, June 27 - July 1, 2006, Bloomington, IN, retrieved March 5, 2008 from [http://garnet.acns.fsu.edu/~abaylor/PDF/ICLS\\_frustration.pdf](http://garnet.acns.fsu.edu/~abaylor/PDF/ICLS_frustration.pdf).

Berscheid, E. & Walster, E. (1974). Physical attractiveness. In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, 7, New York, NY: Academic Press, 158-216.

- Bickmore, T. & Cassell, J. (2005). Social dialogue with embodied conversational agents. In van Kuppevelt, J., Dybkjaer, L. & Bernsen, N. (Eds.), *Advances in Natural, Multimodal Dialogue Systems*, New York, NY: Kluwer Academic / The Netherlands: Springer, 23-54.
- Brahnam, S. (2001). Creating physical personalities for agents with faces: Modelling trait impressions of the face. *Paper presented at the User Modelling 2001 (UM'01) Workshop on Attitudes, Personality and Emotions in User-Adapted Interactions*, July 13-17, 2001, Sonthofen, Germany, retrieved March 23, 2006 from <http://aos2.uniba.it:8080/sonthofen/brahnam-f.pdf>.
- Brewer, M. (1988). A dual process model of impression formation. In Wyer, R. & Srull, T. (Eds.), *Advances in Social Cognition*, 1, Hillsdale, NJ: Erlbaum, 1-36.
- Cook, M. (1979). *Perceiving others – The psychology of Interpersonal Perception*, London, England: Methuen.
- de Rosis, F., Pelachaud, C. & Poggi, I. (2004). Transcultural believability in embodied agents: A matter of consistent adaptation. In Payr, S. & Trappl, R. (Eds.), *Agent Culture: Human-Agent Interaction in a Multicultural World*, Mahwah, NJ: Laurence Erlbaum Associates.
- GameGirlAdvance (2004). Comments: It's time for androgyny, it's just Vaan! *Comment on weblog post at GameGirlAdvance by pseudonym kyoujin*, May 20, 2004, retrieved March 5, 2008 from [http://www.gamegirladvance.com/cgi-bin/mt-comments.cgi?entry\\_id=532](http://www.gamegirladvance.com/cgi-bin/mt-comments.cgi?entry_id=532).
- Graesser, A., Lu, S., Jackson, J., Mitchell, H., Vetnura, M. & Oleny, A. (2004). Autotutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, 36 (2), 180-192.
- Gulz, A. (2005). Social enrichment by virtual characters – differential benefits. *Journal of Computer Assisted Learning*, 21 (6), 405-418.
- Gulz, A. & Haake, M. (2006a). Design of animated pedagogical agents – a look at their look. *International Journal of Human-Computer Studies*, 64 (4), 322-339.
- Gulz, A. & Haake, M. (2006b). Visual design of virtual pedagogical agents: Naturalism versus stylization in static appearance. *Paper presented at the 3rd International Design and Engagability Conference @ NordiCHI 2006 (iDec3)*, October 15, 2006, Oslo, Norway.
- Gulz, A., Ahlner, F. & Haake, M. (2007a). Visual femininity and masculinity in synthetic characters & patterns of affect. In Paiva, A., Prada, R. & Picard, R. (Eds.), *Affective Computing and Intelligent Interaction*, Berlin/Heidelberg, Germany: Springer, 654-665.
- Gulz, A., Haake, M. & Tärning, B. (2007b). Visual gender and its motivational and cognitive effects – a user study. *Lund University Cognitive Studies*, 137.
- Hall, L., Woods, S., Dautenhahn, K., Sobral, D., Paiva, A., Wolke, D. & Newall, L. (2004). Designing emphatic agents: Adults versus kids. In Lester, J., Vicari, R. M. & Paraguacu, F. (Eds.), *Intelligent Tutoring Systems*, Berlin/Heidelberg, Germany: Springer, 604-613.
- Johnson, W. (2003). Interaction tactics for socially intelligent pedagogical agents. *Proceedings of the 8th International Conference on Intelligent User Interfaces (IUI 2003)*, New York, NY: ACM Press, 251-253.
- Kalick, S. (1988). Physical attractiveness as a status cue. *Journal of Experimental Social Psychology*, 24, 469-489.
- Laurel, B. (1993). *Computers as Theatre*, Reading, MA: Addison-Wesley.
- Massaro, D. (2004). A framework for evaluating multimodal integration by humans and a role for embodied conversational agents. *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI 2004)*, New York, NY: ACM Press, 24-31.
- McArthur, L. (1982). Judging a book by its cover: A cognitive analysis of the relationship between physical appearance and stereotyping. In Hastorf, A. & Isen, A. (Eds.) *Cognitive Social Psychology*, New York, NY: Elsevier, 149-209.
- Moreno, R. & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology*, 31, 186-207.

- Moreno, R., Mayer, R., Spires, H. & Lester, J. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-213.
- Moreno, K., Person, N., Adcock, A., Van Eck, R., Jackson, G. & Marineau, J. (2002). Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In *Working Notes of the 2002 AAAI Fall Symposium on Etiquette for Human Computer Work (Technical Report FS-02-02)*, Menlo Park, CA: AAAI Press, retrieved June 11, 2007 from <http://web.media.mit.edu/~bkort/AAAI/Moreno-paper.pdf>.
- Nass, C., Isbister, K. & Lee, E.-J. (2000). Designing and evaluating conversational interfaces with animated characters. In Cassell, J., Sullivan, J., Prevost, S. & Churchill, E. (Eds.), *Embodied Conversational Agents*, Cambridge, MA: MIT Press, 374-402.
- Next Generation (1998). Girl trouble. *Next Generation Magazine*, January 1998, 37, retrieved June 11, 2007 from <http://www.electric-escape.net/node/542?PHPSESSID=50a6bdb50478cae8b6d3424d2c79dd57>.
- Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobrepeze, P., Woods, S., Zoll, C. & Hall, L. (2004). Caring for agents and agents that care: Building emphatic relations with synthetic agents. *Proceedings of the 3rd International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS 2004)*, New York, NY: ACM Press, 194-201.
- Reeves, B. & Nass, C. (1996). *The Media Equation: How People Treat Computers, Televisions and New Media Like Real People and Places*, New York, NY: Cambridge University Press.
- Schleiner, A.-M. (2000). *Mutation.fem*. Text for the “Alien Intelligence” Exhibit, Kiasma, Museum of Contemporary Art, Helsinki, Finland, retrieved June 11, 2007 from <http://www.opensorcery.net/mutation/>.
- Schneider, D. (2003). *The Psychology of Stereotyping*, New York, NY: Guilford Press.
- Smith, E. & Medin, D. (1981). *Categories and Concepts*, Cambridge, MA: Harvard University Press.
- Voelker, D. (1994). *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. Doctoral dissertation, Stanford University, Stanford, CA.

## Communities of Ethical Practice: Using New Technologies for Ethical Dialectical Discourse

**Linda Newman**

Senior Lecturer (Early Childhood), School of Education, University of Western Sydney  
Phone: +61-2-4736 0048 // Fax: +61-2-4736 0400 // Mobile: 0414 376 072 // l.newman@uws.edu.au

**John Findlay**

Director, Zing Technologies, Australian Technology Park, Sydney, Australia,  
Phone: 0414 761 246 // Fax +61-2-9209-4661 // jgf98@uow.edu.au

### ABSTRACT

The authors report on a project in which a new experiential form of professional learning combined ethical thinking processes with a collaborative meeting technology known as the Zing team learning system (ZTLS). A new software program called *Working Wisely* was built by the completion of the project. The ZTLS in combination with *Working Wisely* facilitates communities of practice (Fleer, Edwards, Hammer, Kennedy, Ridgway, Robbins, & Surman, 2006; Wenger, 1998) focused on ethical issues. The participants in the pilot project included 10 staff members of an early childhood education facility (long day care centre) who met for a series of professional development sessions in a learning community that combined industry, university, and commercial stakeholders. By the end of the project, participants were able to move their knowledge from individual tacit knowledge to shared articulated knowledge with an ethical focus. This paper focuses on the facilities that were embedded in the software to develop thinking and discourse processes. The authors faced challenges in translating paper-based thinking processes with numerous steps into simple methods that could be undertaken on an information communication technology (ICT) platform using the ZTLS. These were resolved by the development of miniature and iterative dialectical thinking processes within thinking steps that facilitated discourse. The use of dialectical and ethical discourse in this community of practice led to positive outcomes and has considerable potential in many other areas of academic endeavour. We have called our new type of interaction Ethical Dialectical Discourse.

### Keywords

Professional learning, Ethical discourse, Information communication technology, Team learning, Zing

### Introduction

Human diversity and difference are increasingly acknowledged in literature and valued in contemporary education and care contexts. Now, more than ever, professional educators need to deal with complex issues and work hard to avoid marginalising some people, or to negotiate or resolve misunderstandings. Issues frequently arise at the confluence of different belief systems, lifestyles, sexual orientation, ability, moral positions, and the cultural norms of parents, children, and co-workers (Newman, Mercer, Fleming, & Michail, 2004). Professionals, particularly those working in children's centres, are often unable to easily gain access to professional development courses to learn about contemporary professional and societal changes and expectations. Workshop-based courses delivered by professionals are difficult to arrange for this group because of limited financial support, long working hours, early starting times, late finishing times, shift work, and the isolation of small workplaces. "One-off" staff development sessions are also increasingly recognised as unsustainable. Despite current imperatives in our globalised ICT-based knowledge economy (Newman & Ashton, 2007), many individuals in this population have limited access to e-learning courses because they are inexperienced computer users with basic or no computer skills and have extremely limited access to computers in classrooms, offices, or homes. This situation is exacerbated by limited resources for hardware and software purchase, repair, and upgrading. The population includes many mature-aged women who have not grown up using computers, have little desire to use computers for work or leisure, and at home are the last person to get a turn on the family computer.

### Development of a project for early childhood staff team learning: purpose and procedures

In response to needs identified by Penrith City Council, in outer Western Sydney, NSW, Australia, the authors created a series of 10 staff development workshops to conduct with a child-care centre staff. The purpose was to

more soundly underpin staff interactions with families on ethical practice and to develop their technology skills. We aimed to shine a light on their dealings with parents, but also with each other and the children, with a view to improving ethical practice for sustainable change. In this paper our focus is on outlining the process undertaken, rather than on detailed investigation of results. While we are reporting on a small regional development project, rather than a research project, we feel that our tentative findings lay the groundwork for later research and indicate the potential of the ZTLS in conjunction with the ethical and dialectical discourse that we outline later in the paper.

The ZTLS is a tool designed to support higher-order thinking and human interaction (Findlay, 2003). Fitzgerald and Findlay (2006, p. 2) claim such tools are “no longer merely desirable but now essential if humans are to productively engage with the increasing complexity and uncertainty that arises from accelerating technological and social change.” The ZTLS helps users to take on the role of facilitator or teacher. It consists of a number of computer keyboards (in this project, we used one per participant), one computer, and a projected display. It allows for remote connection of users via the Internet, but in this project participants met face to face. One person in each session was designated facilitator and had control of the software. During the project, facilitation was commenced by the commercial partner (author 2), quickly assumed by the university partner (author 1), and shortly afterwards taken on by several child-care centre staff. Participants each contributed their ideas using their keyboard, following small group discussion, in response to trigger questions or images on the screen. Everyone could see the ideas on the screen as they were being generated, allowing for scaffolding and knowledge building. The facilitator then encouraged participants to pull out the main themes or ideas from each response to generate a collaborative team response. What resulted was a thinking or decision-making journey (Fitzgerald & Findlay, 2006), generated by the team and recorded within the program for future reference and emailing to participants at the end of the session.

The ZTLS concept was designed to create contexts and spaces for thinking and learning and is partially based on the work of Vygotsky (1978), who believed that humans develop through the engagement with tools such as language and signs, which involves not only the transformation of cognitive capacities in the individual but in human society as well. Fitzgerald and Findlay argue that the next generation of learners is becoming more independent of their parents and teachers through the use of advanced tools such as computers, multimedia, and the Internet. They argue that if we are to prepare people for a world in which “work, knowledge and communities are being radically and constantly transformed” (Downes, Fluck, Gibbons, Leonard, Matthews, Oliver, Vickers, & Williamson., 2001, p. 16), then new kinds of tools which are more relational and reflective are needed to give young people the opportunity to access, design, create, facilitate, and improve their own learning activities (Fitzgerald & Findlay, 2006). The theory that informs the ZTLS is in line with current thinking in early childhood education, which draws on post-modern approaches to engage with the complexities and diversities inherent in work with children and families.

The project consisted of ten two-hour workshops. What was covered in each session was somewhat limited by the tiredness of participants by the end of an eight-hour day with children and families followed by the two-hour sessions. This is noted as an issue for the professional development of early childhood practitioners. Workshop 1 took the form of an orientation to the project, familiarisation with the technology, ethics protocols, the completion of a short survey, and an ethics audit of the centre. The survey and audit were repeated at the end of the program. Over subsequent workshops, extensive exploration of personal values was the focus of sessions, and more time was spent on this aspect than was originally anticipated as complex and conflicting views emerged and were discussed and documented. As workshops progressed, the focus shifted to exploration of professional values and the process of documenting the team’s values in relation to children, colleagues, families, the profession, and the community. This was accompanied by an attempt to draft “if-then” statement such as “if we respect children, we respect children by... (insert action).” The intention was to complete a values statement as the foundation for curriculum development. This aspect of the documentation was not able to be completed because extra time had been spent at the start on the personal values aspect of the program, but the centre staff intended to continue with their work after the project finished. Staff were introduced to the Ethical Response Cycle (ERC) (Newman & Pollnitz, 2001), which is a process-based model for the investigation and resolution of ethical issues. By the end of the project, the staff who attended the last workshop had each pledged to work on one aspect identified in the summative ethics audit as in need of attention.

The project brought together two new tools for the first time. The Ethical Response Cycle is a process (available both on paper and digitally) for exploring and resolving ethical dilemmas. It was combined with the computer-based team learning system (ZTLS). The ERC provides a structured thinking approach for the resolution of competing professional, legal, and ethical considerations. It provides a guideline process for thinking through, acting upon, and

documenting issues, problems, or dilemmas that have an ethical component (Newman & Pollnitz, 2002). It contains eight phases that encourage users to consider their dilemma in relation to legal issues, professional issues (codes and policies, etc.), ethical principles (five are suggested), ethical theory (three paradigms are suggested) and “Informed Inclination,” a professional concept introduced in the ERC. The final three phases prompt the user to make a decision, take action, and document the action. The whole process is surrounded by reflection and negotiation. The two tools were used as a springboard for the development of a new software title called *Working Wisely*. This software provides a framework and process for context specific development of values-based curriculum frameworks.

*Working Wisely*, using the ZTLS, supports decision making or learning groups engaged in higher-order thinking and human interaction (Findlay, 2003). It does not rely on the presence of an expert in the topic under discussion as question sequences previously crafted by an expert in the content area provide the triggers for group discussion. Images or provocations (diagrams, photos, etc.) often support the questions. The tool’s embedded etiquette guides the facilitator as participants work through a talk-type-read-review process for each question. Thus, it is a perfect tool in communities of practice where knowledge is shared rather than hoarded away for personal use (Ford, 1997, cited by Kenyon & Hase, 2001). For example, a warm-up question asks participants to nominate their favourite children’s story and asks why. The talk phase of the embedded etiquette often sees participants recalling childhood memories. These memories are triggered by each other to enable participants to recall not only the story, but the context and the feelings involved. The “why” quickly draws out values related to relationships, family, safety, security, etc. When shared with the group, rich meaning around adult–child relationships emerges. A sensemaking (Weick, 1979) step in this procedure helps participants transform their tacit knowledge into prototypical concepts, models, rules, methods, and theories. For example, one module in *Working Wisely* called “two mums” takes participants through a scenario in which a Master’s student, deputy school principal, and press secretary grapple with personal and political issues related to diverse family patterns, such as “What happens when your funding body doesn’t like a video you have made that depicts a successful family as having two mothers or a parent with a severe disability?” Participants are first guided through the steps that allow them to examine the ideas and issues with the help of the ethical theory and processes embedded in the questions and images. They are then asked to look at the process they have undertaken and design their own prototype or process for the resolution of future dilemmas they may encounter.

Embedding the ERC within the ZTLS facilitates meaning-making, deliberating, and theorizing, which are essential elements in both the development of the practical wisdom (Goodfellow, 2003) that underpins Informed Inclination (Newman & Pollnitz, 2001) and ethical dialectical discourse. The delivery method also offers a means of subtly improving computer skills and motivation to use technology. One participant had previously never touched a keyboard, and was quite reluctant to do so at the beginning of the project. Her response in the final evaluation, when asked, “Think back to your first Zing session. How did you feel about the session?” was “PAINFUL.” By the end, however, participants were waiting for her to finish because she was so keen to record her views.

The ZTLS software can be used in place of an accomplished adult to scaffold learners and provide learning support (Vygotsky, 1978; Wood, Bruner, & Ross, 1976), creating rich meaning in ethical conversations. This also occurs via a process of bootstrapping or building upon each other’s ideas. The facilitator encourages participants to identify and extract the main themes, concepts, or ideas from each group of responses to formulate new memetic (Dawkins, 1976) units, which are prototype units of knowledge. The facilitator needs only to know how to set up and operate the computer system and how to follow the ZTLS etiquette. Different learning styles are integrated through an embedded talk-type-read-review process. By following this process, a dynamic and powerful learning community is created. Early in the project, participants were already encouraging each other to draw common themes from their individual responses.

In the *Working Wisely* project, participants discussed each topic in dyads to allow the maximum amount of interaction in the shortest possible time. The group was allocated three to ten minutes to discuss a topic before contributing their own ideas via the keyboard. The workshops comprised sequences of questions ranging from initial exploration of personal values through to development of professional values for the team in a variety of areas. Documentation of frameworks for ethical practice in curriculum planning and implementation resulted. Images acted as stimuli for the group to discuss and generate ideas. Each workshop was prepared a few days prior to delivery following input from participants at the end of the prior session. Feedback was used to design the next question sequence. At the conclusion of the project, the workshop questions and a question set from another project were incorporated into the collaborative *Working Wisely* software (Findlay & Newman, 2005).

## **Our theoretical rationale**

### **Conceptualising ethical dialectical discourse**

Our work was informed and influenced by theoretical perspectives relating to sociocultural learning in communities of practice, ethics, discourse, and professional wisdom. We used a case study approach as it is generally recognised as the most effective way to learn about professional applied ethics (Newman & Pollnitz, 2005). Ethical theory was introduced within a series of scenarios and then applied to real cases to build ethical knowledge and strategies. Ethical dilemmas are used to provide a rich environment in which to explore and resolve complex issues with numerous competing and often equally valid imperatives to consider. Such an approach is a good fit as we move beyond the knowledge age (Findlay, 1997; Findlay, Crawford, & Lee, 2002) to a wisdom age, in which the most important activity is the wise application of knowledge and values (Findlay, Crawford, & Lee, 2002; Lakoff, 2002). Discussion of cases and dilemmas can be enriched as the facilitator guides participants towards higher levels of human interaction.

Case-study methodology was enriched in the development of *Working Wisely* with adult learning approaches that move learners towards the heutagogy, a self-directed approach to learning that underpins our approach (Ashton & Newman, 2006), which supports collaborative team learning. Six kinds of discourse (Jenlink, 2001) also informed the development of *Working Wisely* and partially led to the development of our new concept of ethical dialectical discourse.

### **Creating rich meaning in ethical conversations**

In conceptualising ethical dialectical discourse we chose to build on some existing paradigms. Each paradigm has its benefits and accompanying shortcomings for rich, ethically based conversations in communities of practice. Professionals need the language and the language tools of ethical dialectical discourse to explore, come to terms with, and resolve complex ethical issues within their learning communities. We now examine the components of discourse that we have drawn from to create the concept of ethical dialectical discourse.

## **Discourse**

Jenlink (2001) describes six kinds of discourse:

- (1) *Monologue*, which is one-way communication and rule-following (Isaacs & Senge, 1999), where one person speaks and others may listen with little or no counter-influences of ideas.
- (2) *Discussion*, which is rule-revealing (Isaacs & Senge, 1999), where views are exchanged in an adversarial way with little or no interpenetration of ideas.
- (3) *Dialogue*, which can be of two types, reflective or generative, the former being rule-reflecting (Isaacs & Senge, 1999) and the latter rule-creating, in which the participants are open to each other and suspend judgment, beliefs, and assumptions in the interest of creating new knowledge that simultaneously serves the group and individual interests.
- (4) *Dialectical*, originally considered to be disciplined inquiry using rational argument sometimes presented as truth and which is equally rule-revealing and adversarial. Dialectical discourse has also been interpreted by others as transformational and integrative (Engestrom, 1987), where the differences are resolved into a new and more complete idea or concept about which the participants in the discourse may assign new meaning.
- (5) *Ethical* (Banathy, 1996), which is controlled by social rules — values, morals, and ethics — but which interacts at the boundaries of complex cultural systems and can also be a tool for the wise application of knowledge.
- (6) *Post-formal* (Horn, 1999), which could also be called transcendental, and is a discussion about power and the consequences of the discourse in which we are engaged, and which is shaped by patterns, processes, and contexts. McLean describes a similar level of discourse as social reconstructivist (1999).

## **Ethical dialectical discourse**

Professional ethical wisdom is underpinned by informed discourse, reflection at the social reconstructivist/ethical level (McLean, 1999), and notions of transcendental discourse (Horn, 1999). The discourse model employed by the



*Working Wisely* program in the ZTLS context employs elements of dialogical, dialectical, ethical, and transcendental discourse to develop ethical wisdom informed by ethical dialectical discourse.

Ethical dialectical discourse moves beyond any of the aforementioned individual discourse concepts because it

- is disciplined inquiry
- uses rational argument but is not adversarial
- is rule-creating
- is reflective, generative, and integrative
- allows participants to be open to each other and suspend assumption and judgment
- considers values, morals, and ethics
- considers complex cultural systems
- considers power and consequence, and
- is socially and cognitively transformative.

We believe that ethical dialectical discourse contains the potential for transformation and for the application of professional wisdom in complex times.

## **Developing the team learning system guided questioning process: our challenges**

During the course of the project the authors, particularly author 1, faced a number of challenges in translating accustomed thinking about the preparation of traditional workshops and professional ethics content into a collaborative computer-based inquiry environment that incorporated ethical dialectical discourse. Talking about ethics is abstract at any time, and some thinking processes are difficult to capture as pre-defined questions. The facilitation method required to engage in thinking in this manner was new to author 1, and the ethics content was new to author 2. If the authors of the program were to create a successful software program to move users from a basic understanding of professional ethics to a more sophisticated, analytical, ethical, and dialectical method of higher-order thinking, their own thinking required adaptation too.

The ZTLS system presents open-ended questions that guide the participants through a pre-determined thinking journey that emulates the kind of thinking that a domain expert would undertake. Six main kinds of question processes were employed to facilitate ethical dialectical discourse in this project. We have adopted the following nomenclature to reflect the purpose of the level of question and the guided questioning process: orientation, sequential, extended sequential, self-revealing, conditional, and integrative questions.

### **Orientation**

The orientation questions are presented when a group is using the ZTLS for the first time. The way in which the first three questions are presented is critical to the successful use of the tool. The first question allows participants to find where their responses will be located on the screen, in which writing space their entries will appear, and how to use their keyboard to edit ideas or submit them to the common space. Following are orientation questions from Workshop 1 (*Working with the Joys and Challenges*):

- Type anything you like: the words of your favourite song, a list of what you had for breakfast, or “the quick brown fox jumps over the lazy dog.”
- What is your favourite children’s story and why?
- Describe the nicest/most fantastic/rewarding moment(s) in your work with children or in children’s services.

When the second and third questions are presented, the facilitator models the ZTLS talk-type-read-review etiquette, sharing ideas, reading ideas aloud, and seeking out the underlying concepts and themes from the ideas of the group. The etiquette has a similar function to the think-pair-share method (Lyman, 1987) used in non-ZTLS settings, which requires participants to think alone first, then discuss with a neighbour before sharing their ideas with other participants who both see and hear all responses. This allows for auditory, visual, and kinaesthetic learning. Such methods orchestrate the way participants interact with each other and contribute to the overall “dialectical” integration in an iterative manner. The orientation questions also serve as a way of helping the group develop a level

of comfort with each other and quickly move participants from the general (Question 1) towards the professional and values based (Questions 2 & 3). There is also a team formation effect (Findlay, 1997).

## Sequential

The sequential questions have a simple logic that requires little facilitator intervention. The facilitator merely follows the talk-type-read-review etiquette for each question. This type of minimal intervention by a facilitator, sometimes known as servant leadership (Jackson, 1986), is the highest form of facilitation because it transfers responsibility for the activity to the participants, who take ownership because they believe they have achieved everything by themselves (Heider, 1985). Scaffolds within each question are used to direct attention, especially in situations where participants may have had difficulty understanding the question or are unable or unlikely to generate ideas without some assistance. Scaffolds may also be used to encourage responses to fit a particular format. Usually the questions are arranged in a specific order, which allows participants to collect data about a topic and, via the thinking process, convert that data into some higher form such as a concept, mind map, check list, action plan, or decision. These questions are from Workshop 3 (the Shoulds and the Should Nots):

- Think back to something you did as a child or a teenager that you are proud of. In 25 words, without necessarily describing the specific incident, write about why you were proud and what made you do what you did.
- Think back to something you did as a child or a teenager that you are not proud of. In 25 words, without necessarily describing the specific incident, write about what was not right, fair, honest, kind, caring, or just about what you did and why you think you did it.

The inclusion of words like “fair,” “honest,” and “just” flags the beginning of the embedded introduction of ethical theory and begins to introduce the language and prototypes for ethical decision-making.



*Figure 1. Automatic Teller Machine Fairy image*

## Extended sequential

The extended sequential questions often incorporate an image as a stimulus. This type of questioning approach was applied to a series of dilemmas that allowed the participants to experience the dynamics of a hypothetical situation. A rich description of each stage of the dilemma is presented as an image (see Fig. 1) attached to a short version of the

question that asks each participant to describe what he/she would do under the circumstances. The dilemma sequence becomes more complex as the workshop proceeds. These questions are from Workshop 2 (Adhering to the Unenforceable):

- The “Automatic Teller Fairy” has been helping out many people in your town for a week now. Word has spread. The error has been discovered and rectified. The Daily Bulletin reports the names of everyone who used the teller and how often. Your mother rings that night to say she saw your name in the paper. What do you say and why?
- How do you feel when your employer raises the issue the next day, and why?

## Self-revealing

A third type of question used in the workshop, known as self-revealing, was designed to challenge and help participants to examine their own instinctive patterns of behaviour and discover their own values and own biases in interacting with people from backgrounds different from their own or who have different beliefs or values. Participants viewed an image (see Figure. 2) and then described what they saw. The image was displayed a second time, this time with a description of what was really happening. Figure 2 shows one of the images, a photograph of the first author and a friend sitting on the floor eating food with their hands in India. The following sequence is from Workshop 8, What Values are Represented Here?

- Write about what you can see in this picture.
- This is what was really happening.
- List some of your personal values including your ideals, principles, qualities, accomplishments, actions, and circumstances.

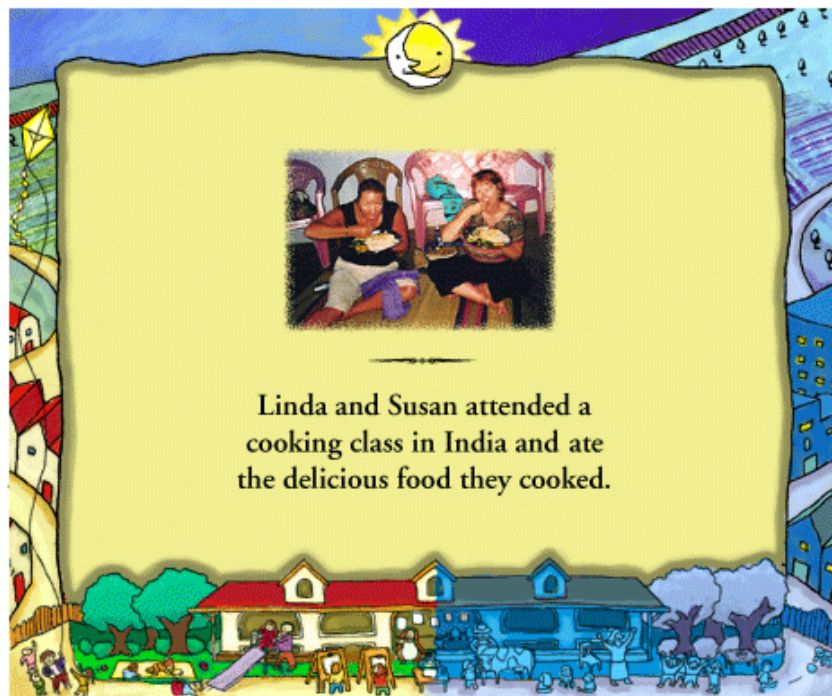


Figure 2. Values Session image

This workshop arose following a conversation among the project members about the appropriateness of allowing children in the centre (who were from a variety of cultural backgrounds) to eat food with their fingers.

## Conditional

The conditional questioning process requires participants to collect information which then becomes the subject of the next thinking process. This was one of two processes which proved challenging for the software developers to

capture in a way that was easy for a facilitator to routinely follow. The software allows the facilitator to add a thinking process at any time and attach this to any idea that has been generated. The facilitator can also “drill down” to explore an issue in more or minute detail.

The purpose of this process was to examine personal and professional values with a view to developing a context-specific values-based curriculum framework for the early childhood centre. The participants decided who the important stakeholders in their working lives were. Then they decided what was important in relation to each group in order to define the value base of their early childhood service. They then linked desired actions to the values chosen. The stakeholders identified by the group were staff, children, families, the community, and the early childhood profession. The aim was to help bridge the common dissonance between espoused values and values in action.

Workshop 9, (Working with Stakeholders), began with participants brainstorming a list of stakeholders of the child-care centre followed by a list of values to guide interactions with each stakeholder as stakeholders would differ for each group of users. Following this, for each value selected, participants were asked (e.g., for the stakeholder group “children”):

- Explain why we should have this value? e.g., for respect, because children are citizens in their own right.
- How would we use this value? e.g., for respect, ask children what they think.

One purpose of this elaboration was to encourage participants to articulate their team values and resultant practice. The documented values can be used as the basis of communication with families, as proactive planning for ethical practice, and as a service evaluation process. A prototype was built that allowed participants to articulate “if I believe *this*, then I will do *this*” statements.

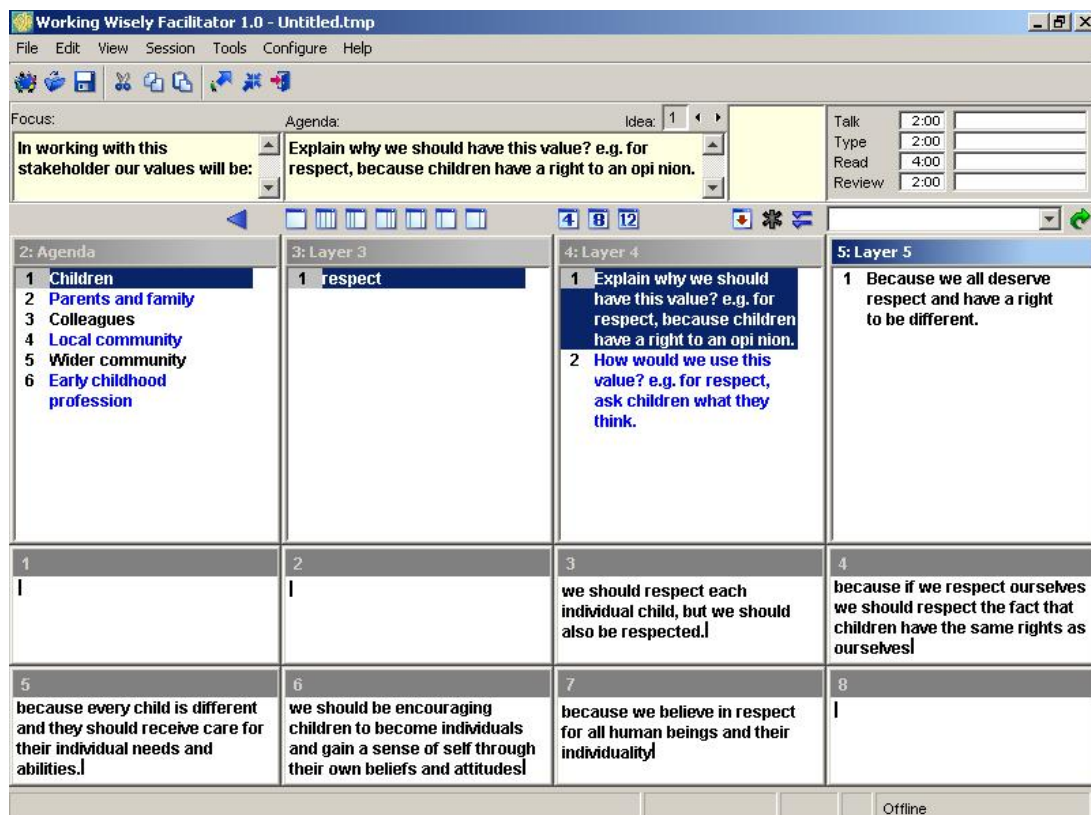


Figure 3. Drilling down four layers to explore values related to stakeholders

Such a conditional process that drilled down four layers could not be incorporated easily into a logical, sequential thinking process or prepared in advance and applied to all stakeholders to be identified. The solution found by the authors was to present the entire process across four layers using the multiple layers mode of the software. In this way, each stage of the brainstorm was made visible, including the list of stakeholders identified in the session and the

values related to each, accompanied by the linked actions (See Figure 3), in which the lower eight boxes are participant playspaces where ideas are created. Ideas accumulate in the upper space, the teamspace. Here the teamspace is divided into four layers to allow the facilitator to drill down to explore the relationship among stakeholders, a value, and the reason for holding that value.

## **Integrative**

The Ethical Response Cycle, which is at the heart of *Working Wisely*, is an integrative questioning process that requires participants to engage in a dialectical or ethical social re-constructivist style of discourse that we have now named ethical dialectical discourse. The ERC process results in the creation of a unique new idea or position through iterative integrative thinking informed by ethical considerations.

When the authors first attempted to translate the original ERC into a ZTLS sequence, they encountered unanticipated difficulties. The problem was that the structure of many thinking processes is not obvious and, in sessions not supported by technology, flexibility is achieved by expert facilitators switching modes and taking shortcuts with the process, often in response to on-the-spot nonverbal cues from participants. If the facilitator does not explain the change in the focus of attention to the group, the participants can become disoriented and uncertain of what to do. Most facilitators do this instinctively and would not be able to articulate their methodology. In the computer-supported facilitations, the change of mode became much more obvious, and we needed to deconstruct the process to create a solution.

When fully documented for *Working Wisely* development, the ERC was found to involve eight phases (the law, professional considerations, ethical principals, ethical theories, Informed Inclination and a judgment, action and documentation). These eight phases then contain 26 separate steps. If followed assiduously in a software program, rather than in its original form as a thinking process, this would have been de-motivating for participants and so time-consuming that they would be unlikely to revisit the methodology.

In response to the authors' own dilemma about how to capture the process in an effective, engaging, time-efficient and yet still thorough manner, we devised integrated questions that we feel still represent the ERC with integrity. The revised process has only eight steps, but within each step is a series of sub-processes that require the participant to engage in a miniature thinking process, before submitting a response. Each step is accompanied by an image with a check list of things to think about at that stage of the process.

Through engagement with the process of the images and thinking steps, participants undertaking ethical dialectical discourse in the community of ethical practice can collaborate on the resolution of current, contextually based issues.

## **Some snapshots of the Zing *Working Wisely* project**

### **Values and knowledge**

The focus of this paper is on the development of the *Working Wisely* platform rather than on reporting full findings about participants' development of ethical practice; however, it is limiting to discuss one without the other so some examples are also presented here. Because of the small number of participants, we make no claims for statistical reliability. Our initial feelings from this small pilot project were that the marriage of the ZTLS with experiential ethical dilemmas and a robust dilemma resolution process was an engaging and successful way to learn about ethics and apply ethics in a meaningful way. The focus of evaluation was on the ethics content, rather than on the ICT tool. However one participant, in evaluating the workshops, said, "The technology of Zing and its use was great — very interesting. The discussions and self-awareness of our own belief systems were eye-opening." Another said "... The technology was explained well and the system was reasonably easy to use." People in the project who had never previously used computers were able to easily use the system, and staff quickly stepped forward to take on the role of facilitator. Belinda said, "I really enjoyed attending and 'facilitating' these workshops. I would enjoy imparting my new knowledge with other groups." Participants found that they acquired an increased understanding of their roles as ethical professionals while participating in entertaining or real-life play with situations or issues.



When used well, the tool supported ethical dialectical discourse, although at times the groups reverted to other less robust forms of discourse such as discussion or monologue, which can have a lesser effect on the knowledge creation process if the facilitator is not careful to maintain momentum and control over the process. One participant reported that “some [sessions were] a little boring and others made you think.” This participant missed several sessions and seemed less involved than others. A follow-up interview would have been useful but was not possible at the time. Reclaiming robust discourse when participants became disengaged was achieved by a counter-intuitive method of minimalist but clear and well-executed interventions embedded in the ZTLS etiquette. This was timed to provide scaffolding for the learners as required.

Throughout the project, we needed to be aware of limitations. Initially, in the project described, the dyadic discussion model embedded in the tool’s etiquette was resisted by participants. They seemed initially more comfortable to share ideas with a chosen friend/colleague than with an immediate “neighbour” around the table. They preferred, overall, to quickly type in their own ideas. Participants did however accept monologue or group discussion with turn taking as a way of conversing. Where views were expressed by an authority/power figure (director or project leader), there was initially a tendency by some to copy the responses. It is also possible that participants did not respond publicly but may well have entertained their own views privately. In the project evaluation, one participant summed this up in response to a question about what was learnt. She said “not to copy Jan” (the director). Interestingly, it had been noted by the project leader that the director began to slow her entry of typed responses, presumably to allow her staff to form and express their own ideas in response to noticing some participants “copying Jan.” She had previously expressed a genuine desire to the project leader for the team to express their tacit knowledge and build team understandings.

After several rounds of question asking, participants began to adopt the etiquette and were able to engage in higher-level discourse automatically. Staff became genuinely interested in what their colleagues thought or felt about issues, especially if those ideas resonated with or complemented their own or offered a perspective they had not previously considered. Belinda said “[The sessions] were great. The small groups and interactions were very informative as quite often these issues are not ‘pulled apart,’ just discussed. They made me think and relate it to my own practice.” The workshop experience was a fertile ground for the development of practical wisdom (Goodfellow, 2003) as participants drew from their professional experience and built upon previous knowledge and practice to create new knowledge. “PA system” (some participants used pseudonyms) said, “I think the workshops were great and made people more aware of how different people can interpret the same thing in different ways.”

Instances of dialectical discourse (Isaacs & Senge, 1999) were evident and captured within the programme when the participants sought to resolve the conflict between their competing stances and build new knowledge. Also evident were examples of ethical discourse in which participants developed new understandings across the boundaries of complex cultural systems and social rules (Banathy, 1996). Post-formal or transcendental discourse (Horn, 1999) was evident in consideration of how participants’ discourse would impact others (for example, children and families). In response to the questions about the values the participants felt should be espoused by the staff of the centre, for example, participants cited “respect,” “nurturing,” “celebrating differences,” “freedom of expression and acceptance of differing levels of achievement,” and “caring relationships.” The participants then explored how each value would be implemented. For the “respect” value, the participants expressed a much richer variety of actionable activities than they did when they embarked on the program nine months earlier. In summarizing the responses of the group, participants said they would celebrate differences by listening to children; be aware of the child and his/her family and respond to the differences; act on children’s interests and capabilities; help children feel good about themselves and others; ask their families what they think is important for their child to learn about and do; and plan for each child’s individual needs, strengths, and interests in a flexible and spontaneous environment. They also provided justification for their actions, such as the following comments: “because every child is an individual and we should cater to their needs and abilities,” “to gain a sense of self through their own beliefs and attitudes,” “because we all deserve respect and have a right to be different,” and “children should learn about respect.”

In the evaluation of the program, the participants gave examples of how they had changed during the activity and had developed as professionals. One participant said, “The [ERC] was a good learning tool — particularly for those who are impulsive in decision-making.” Another said, “[The project] has allowed [me] to question and consider my own [values?] and strengthen some in particular. It has enabled me to grow professionally, and this in turn will help in the decision[s] I make with children, families, staff, and life.” Participants moved towards a higher level of reflection (social reconstructivist/ethical) (McLean, 1999) as they engaged and built their Informed Inclination based on new professional knowledge and added experience (Newman & Pollnitz, 2001). Although we did not set out to research

the use of ethical dialectical discourse in this project, we are encouraged to think that it is possible within communities of practice, such as the one we experienced. Research in this area with larger numbers of participants would allow us to confirm or re-consider our inclinations. Research around ethical dialectical discourse using different methodologies would also be useful.

During the course of the project, the styles of response from the participants became more robust and more considered, and exhibited greater explanatory depth. Participants who were initially hesitant or unengaged displayed motivation and interest and grew in confidence, not only in their newfound ability to discuss a complex issue, but also in their ability to express their own divergent opinions. Earlier in the project, they were unwilling or unable to express their ideas in such an articulate manner. They began to use the language of an early childhood professional discussing ethical issues rather than the previous practice of taking a day-to-day superficial and lay, or personal, views and explanations for events that had happened. For example, “when planning for the centre we look at individual needs taking into account individuals’ requirements and beliefs. Each child has an individual profile, menus are planned for each child’s needs, [and] parents’ beliefs and values are taken into [account].”

Through the use of concepts that were for them initially new and lacking in meaning, the participants began to articulate their professionalism. One member focused on her own increased feeling of professionalism and her renewed appreciation of her profession in her evaluative reflections, reporting verbally that she now “felt much more like a professional than a child-minder.” They became more willing to integrate the ideas that they had heard or saw others express and thus moved from their own tacit knowledge to a more academic, theoretically based type of knowledge that was shared across the group. They were, in effect, engaging in a process of knowledge creation. As a post-script, in 2008, two of the participants embarked on tertiary study.

During the project another noticeable feature was the growth in cultural awareness, both of the group learning culture and of issues to do with the cultural backgrounds of colleagues, children, and families. Staff were observed influencing others to change their behaviour if they did not fully engage with the process or regard the content seriously. People became more aware of their own biases including their previously personal, superficial, and simplistic approach to dealing with complex issues. As one participant importantly noted in her evaluation, she wasn’t sure whether she had changed much (others noted changes in her) but she was now aware that her opinions were personal and that she did have biases. This was a huge shift from a desire to impose “normality” onto others from her position of power. At the end of the project, in response to a question about what “ethics is,” the three main themes to emerge were “doing the right thing,” “beliefs and morals,” and “decision making.” There was a noticeable shift in participant responses as to why ethics is important in their work. Initially, responses were varied, including “ethics helps us make decisions in the work of children’s development” and “ethics expresses our morals.” Post project, responses clustered into two themes related to making “good,” “fair,” or “best” decisions and the awareness that decisions affect a number of people, especially young children. These responses reflect a deeper and broader understanding of professional ethics.

Two participants stopped coming to workshops and it was not possible to find out why, although the director did feel that it was for personal family reasons rather than having anything to do with the project or the methodology.

### **ICT dispositions and skills**

The project resulted in the transfer of facilitator skills to the first author (university partner) and subsequently to some members of the child-care centre staff. The second author (commercial partner) began the process of transfer by informing participants of his intention to achieve some kind of skill development during the project. At the start of the first session, the second author informed the group that anyone who wished to become the facilitator should observe what he was doing and saying to conduct the session and to use that as a model for when she became a facilitator. He also let the participants know that the first author would be the first person to be trained in how to facilitate the sessions and that all the participants were expected to help remind her of the mouse and key strokes required as well as the instructions needed and whether or not these instructions should be made clearer. This approach to learning put the university partner and the team on the same level with the technology, taking away some of her “expert” and power status. It helped to create a community of practice, rather than the traditional, expert-led professional development project. The first author assumed facilitation responsibility by session two and, after two sessions as facilitator, handed over responsibility for the technical facilitation to one of the child-care centre staff

who volunteered her interest in performing the role. Neither the first author or staff members had any prior experience with the technology nor were they highly ICT literate at the time.

Another positive change in participants' attitudes towards the project and skill development became evident. Members of the staff were initially unable or unwilling to take responsibility for setting up the meeting room or ZTLS technology in advance to allow a timely start to sessions. They were understandably tired after an eight-hour child-care day. At the beginning of the project, the training room had not been cleared and prepared prior to the arrival of the project leader (first author). At the start of the project, most participants lacked the motivation to undertake the program and saw it as an avenue to achieving some time off from their regular duties. A data projector had to be acquired from the Council head office each time the system was to be used. A computer had to be borrowed as the centre did not have a laptop computer for use. Sessions were sometimes held a month apart and at other times a week apart, depending on the centre's circumstances and what was to be done between sessions. Between sessions some equipment vital to the proper functioning of the system went missing or had to be obtained from the head office. As the project progressed, it was pleasing to see participants keen to begin sessions, unwilling to cancel sessions, keen to prepare the room for sessions, willingly and independently collecting the projector from Council headquarters, and setting up the equipment. They were also capable of running sessions independently by the end of the project.

## Conclusions

Although the strength of our findings is limited by the scope of this small project, we are encouraged to continue our work in this area with some formal research in the future. The integration of the Zing collaborative learning technology and the Ethical Response Cycle method has culminated in a synthesis of the thinking processes involved in ethical social re-constructivist discourse, dialectic, and dialogue, which we have called ethical dialectical discourse. The authors needed to adapt their working habits to devise a workable and useable software program and, in so doing, they devised a methodology that has the potential to facilitate social and professional transformations. The second author, a technology developer, was challenged by the demands of embedding the complex and grey (as opposed to black and white) thinking that ethical discourse involves into a software program. The first author, an academic, was challenged by the demands of adapting her accustomed pedagogy and theoretical approach to an ICT platform. However, we concluded that it is possible to do both. The result was warmly and enthusiastically received by the participants who reported learning a lot, being challenged, and having fun! Leadership emerged in some participants who left the project with a new zeal to engage with both ethics and ICTs. The potential for the ZTLS and *Working Wisely* to embed ethical professionalism into daily practice is exciting and potentially transformative.

Similar approaches could be applied in any corporate, government, or community setting where the issues are complex with many competing cultural interests or where rational decision-making models have proved cumbersome or ineffective. This approach is an important bridge between the tacit or common-sense knowledge usually found in workplace settings among both skilled and unskilled workers and the scientific or theoretical knowledge of professionals and academics. Important to users, the approach is unthreatening, flexible, sustainable, and replicable. It allows professionals to take ownership of their professional development and knowledge creation needs as they practise professional autonomy in communities of practice to explore and resolve "real" and situated issues, rather than engage in remote and theoretical staff development exercises. This allows people to value and articulate their own professionalism and transform their practice.

## References

- Ashton, J. & Newman, L. (2006). An unfinished symphony: 21<sup>st</sup> century teacher education using knowledge creating heuristics. *British Journal of Educational Technology*, 37 (6), 837–852
- Banathy, B. H. (1996). *Designing social systems in a changing world: A journey toward a creating society*, New York, NY: Plenum Press.
- Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
- Downes, T., Fluck, A., Gibbons, P., Leonard, R., Matthews, C., Oliver, R. Vickers, M. & Williamson. (2001). *Making better connections*. Canberra: Commonwealth Department of Education, Science and Training.



- Engestrom, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research, Helsinki: Orienta-Konsultit.
- Findlay, J. (1997). Accelerated Knowledge Creation and Teamwork. *Technology Business Review*, 1, 9–13.
- Findlay, J. (2003). Knowledge creation technologies (KCT). Support for the cultural transformation of schools. In C. Crawford, D. Willis, R. Carlsen, I. Gibson, R. McFerrin, J. Price, & R. Weber. (Eds), *Proceedings of the Society for Information Technology and Teacher Education International Conference 2003*, Chesapeake, AACE, 925-932.
- Findlay, J., Crawford, K., & Lee, M. (2002). Wisdom tools: Hyper-cognitive tools to save us from a small planet. *Proc of the 5<sup>th</sup> Congress of the International Society for Cultural Research and Activity Theory*. Amsterdam. ISCAR.
- Findlay, J. & Newman, L. (2005). *Working Wisely*, Sydney: Knowledge Creation Press.
- Fitzgerald, R. N., & Findlay, J. (2006). *Team learning systems as a collaborative technology for rapid knowledge creation* [White paper]. Canberra, Australia: Learning Communities Research Area, University of Canberra.
- Fleer, M., Edwards, S., Hammer, M., Kennedy, A., Ridgway, A., Robbins, J., & Surman, L. (2006). *Early childhood learning communities: Sociocultural research in practice*. Sydney: Pearson Education Australia.
- Goodfellow, J. (2003). Practical wisdom in professional practice: The person in the process. *Contemporary Issues in Early Childhood*, 4(1), 48–63.
- Heider, J. (1985). *The Tao of leadership: Lao Tzu's Tao Te Ching Adapted for a New Age*, Malaysia: Eastern Dragon Books.
- Horn, R. A. (1999). The dissociative nature of educational change. In *The post-formal reader: Cognition and education* (pp. 351–377). Kincheloe, J. L., Steinber, S. R., Hinchey, P. H. (Eds.). New York, NY: Falmer Press.
- Isaacs, W., & Senge, P. (1999). *Dialogue and the art of thinking together: A pioneering approach to communicating in business and in life*. New York: Currency.
- Jackson, G. (1986). *Management of strategic directions for the Sydney Water Board*. Sydney: Sydney Water Board
- Jenlink, P. (2001). Activity theory and the design of educational systems: Examining the mediational importance of conversation. *Systems Research and Behavioural Science*, 18(4), 345–359.
- Kenyon, C. & Hase, S. (2001). Moving from andragogy to heutagogy in vocational education. Retrieved July11, 2008, from [http://www.avetra.org.au/abstracts\\_and\\_papers\\_2001/Hase-Kenyon\\_full.pdf](http://www.avetra.org.au/abstracts_and_papers_2001/Hase-Kenyon_full.pdf).
- Lakoff, G. (2002). *Moral politics: How liberals and conservatives think*. London: University Of Chicago Press.
- Lyman, F. (1987). Think-pair-share: An expanding teaching technique. *MAA-CIE Cooperative News*, 1, 1–2.
- McLean, S. (1999). Becoming a teacher: The person in the process. In R. Lipka and T. Brinthaupt (Eds.), *The Role of Self in Teacher Development*. New York, NY: State University of New York Press.
- Newman, L. & Ashton, J. (in press). Early childhood responses to globalised learning: Flexible and blended delivery in communities of practice. *Australian Research in Early Childhood Education*.
- Newman, L., Mercer, B., Fleming, J., & Michail, S. (2004). Developing ethically based reciprocal interactions in the 21st Century. *Proc of Association of Professional and Applied Ethics Annual Conference*, Univ of Western Sydney, Australia.
- Newman L. & Pollnitz, L. (2001). Will my response be ethical? A reflective process to guide the practice of early childhood students and professionals. [Multi-media kit]. Melbourne: Video Education Australia.
- Newman, L. & Pollnitz, L. (2002). *Ethics in action: Introducing the Ethical Response Cycle*. Watson ACT: Australian Early Childhood Association.
- Newman, L. & Pollnitz, L. (2005). *Working with children and families: professional, legal and ethical issues*. Frenchs Forest, Australia: Pearson Education Australia.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: HUP.
- Weick, K. E. (1979). *The social psychology of organizing*, Reading, MA: Addison-Wesley.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge Univ Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100.

# Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda

**Tel Amiel**

LANTEC, School of Education, State University of Campinas (UNICAMP), Brazil // [tamiel@unicamp.br](mailto:tamiel@unicamp.br)

**Thomas C. Reeves**

Dept. of Educational Psychology and Instructional Technology, College of Education, The University of Georgia, USA // [treeves@uga.edu](mailto:treeves@uga.edu)

## ABSTRACT

The role of educational technologies in improving educational practices and outcomes has been criticized as over-hyped and insignificant. With few exceptions, the state of education has changed less than expected as a result of tools such as computers and the Internet. To a considerable degree, this is due to the minor role educational technology research plays in transforming the use of technological tools in the classroom. This article presents an analysis of technology as a process and as a value-laden system, both of which have substantial consequences to our approach to research. It is argued in the article that design-based research can address some of the deficiencies of other research methods in investigating the role of tools and techniques in the classroom. Through more democratic research practices and recognizing technology as a system beyond its tools, researchers can increase their impact on educational practice.

## Keywords

Educational technology, Design-based research, Values, Methodology, Critique

“It is one thing to state the chief aim of education ... it is quite another thing to pursue this aim in a world which denies the principles on which it rests” (Jeffreys, 1955, p. 13).

## Introduction

New communication, media, and computing technologies have long tantalized educators, policy-makers, and educational technologists as to their prospects for enhancing educational outcomes (Saettler, 1990). Numerous tools ranging from Edison’s film projector through Berners-Lee’s World Wide Web were originally invented for purposes other than education, but they were quickly promoted by educational technologists and others as having enormous promise for enhancing the impact of teaching and learning. Devices now considered to be simple and omnipresent in educational settings were once considered revolutionary and capable of mending social inequity and changing the face of education. For example, in the 1970s, access to handheld calculators was considered to be crucial to raising test scores for underachieving math students, and accordingly math educators and educational technologists led efforts to get calculators into the hands of children learning mathematics. However, once the access gap was closed, the results were found to be much lower than promised (Loveless & Diperna, 2000).

Educational technologies are often viewed not only as solutions to real or perceived inadequacies of traditional instruction, but also as tools for reducing the inequities in educational opportunities around the world. Light (2001) described the rhetoric of social inclusion often associated with new technologies such as cable television. Cable was promoted as a technology to improve not only educational opportunities, but also general access to information. Inequitable access in terms of race and wealth prompted policy-makers and researchers to push for equitable distribution of cable access. Clearly, cable did not achieve educational equity nor increase access to reliable and valid information, but instead is primarily used as increased bandwidth for media outlets. Much of the excitement regarding student achievement that might have been derived from cable in the classroom and other new technologies has faded into disappointment. This trend is nowhere more prevalent than with respect to today’s most heavily promoted technological solution to educational problems, the Internet.

The Internet is in danger of becoming yet another example of society’s all-too-frequent, but usually failed, infatuation with the educational potential of new technologies. Past research has shown us time and time again that, despite all the rhetoric to the contrary, educational technologies do not guarantee big leaps in educational

achievement by any measures, nor have they eliminated the inequitable distribution of learning opportunities (Cuban, 1986). However, the Internet as an educational technology can serve a much more noble and principled purpose. A new approach to educational technology research, one grounded in the *ends* of technology, directed by values and principles, must be pursued.

We argue that traditional predictive research in educational technologies has had limited impact in informing actual use. In other words, educational technology research aimed at examining the influence of tools in the educational process has offered little systematic advice to the practitioner. We argue that recognizing technology as a *process* has implications for how educational technologists conduct research. Once recognized as a process, the aims/ends of technology come to the foreground. We argue that design-based research provides an innovative proposal for research on innovation and education.

## Investment and use

Governments around the world have implemented policies and made substantial funds available to deploy Internet-enabled computers in schools. The cost and maintenance of computers and online technologies in schools far exceed investments in previous technologies.

A dozen years ago, when the World Wide Web was in its infancy and other countries were still experimenting with the computer as an instructional device, the United States already had approximately 5.8 million computers in schools (Office of Technology Assessment, 1995). This trend has not faded. While Oppenheimer (2003), a journalist, calculated a 70 billion dollar “investment” in school technology, even educational researchers have estimated that more than 40 billion dollars have been spent on educational technology infrastructure and training in the past ten years (Dickard, 2003). Regardless of the actual amount, the costs are clearly enormous. Oppenheimer argued that this money should have been spent hiring 170,000 new teachers. Similar investments have been made in higher education institutions with few demonstrable benefits (Hersh & Merrow, 2005; Postman, 2003).

There is little sign that expenditures for school and campus computers are slowing down. The 2005 budget (including discretionary and mandatory appropriations) for the US Department of Education was forecasted at over 70 billion dollars, and almost 500 million were dedicated to state educational technology grants supporting technology integration into the schools (Department of Education, 2005). The same pattern of investment occurs around the world. With the aim of enhancing information and communication technology infrastructure in its schools, England committed approximately 11,200 and 65,000 pounds to each primary and secondary school, respectively, in 2003 alone. As a result, 99 percent of British schools are connected to the Internet (Department for Education and Skills, 2003). Poorer nations have followed suit. Brazil deployed over 53,000 computers in over 4,600 schools around the country as part of a federal government program for technology integration (Departamento de Informática na Educação a Distância, 2002). As demonstrated by the example of these and other countries, technology integration into education is a massive global trend.

What evidence exists that the expenditures on educational technologies such as computers and Internet access have been worthwhile? Although virtually all schools in the United States now have Internet access, recent reports of the use of computing technologies in the classroom (Cuban, 1986; Oppenheimer, 1997, 2003; see also Salomon, 2002) only reiterate what early accounts (Office of Technology Assessment, 1995) have demonstrated: educational technology has been oversold and is generally underused in classrooms around the country. There is no clear evidence of increased achievement resulting from Internet applications in education, nor

[...] has a technological revolution in teaching and learning occurred in the vast majority of American classrooms. Teachers have been infrequent and limited users of the new technologies for classroom instruction. If anything, in the midst of the swift spread of computers and the Internet to all facets of American life, “e-learning” in public schools has turned out to be word processing and Internet searches (Cuban, 2001, p. 178).

Why have we been so naïve in investing this much research time and public money in wiring the schools when past educational innovations, including films, instructional television, and programmed instruction have failed (Cuban,

1986)? What is it about the World Wide Web and other Internet technologies that entice us into believing that, this time, things will be different? It could be that the Internet as an educational delivery system is simply following the hype cycle, that is, over-enthusiasm followed by sharp disappointment with new technological tools. This cycle (found in business and everyday life as well as in education) begins with a peak of inflated expectations leading to disillusionment and finally to a plateau or realistic application. Perhaps we are simply in the midst of a period of inflated expectations that will ultimately lead to more grounded expectations (Gartner Inc., 2004; Rescher, 1980).

However, although skeptics and critics exist, many still believe in the power of the Internet and computers to change the way we teach and learn. Some believe that the online technologies can be used to foster pedagogical change, such as those who promote the constructivist pedagogy movement (see Jonassen, 1991, 2003). At the other end of the pedagogical continuum, extremists claim that the answer to the crisis in education can be solved by employing computers as tutors, without human intermediaries (Bennett, 1996; Jones, 1996). This quote from Lewis Perelman's 1992 book entitled *School's Out* exemplifies the extreme perspective:

Because of the pervasive and potent impact of HL (hyperlearning) technology, we now are experiencing the turbulent advent of an economic and social transformation more profound than the industrial revolution ... In the wake of the HL revolution, the technology called "school" and the social institution commonly thought of as "education" will be as obsolete and ultimately extinct as the dinosaurs (p. 50).

Wang and Reeves (2003) point out that many educators, as well as people in the general public, believe that computers and the Internet are simply much more powerful educational tools and cannot be compared to previous "new" technologies such as the television. Although large-scale success stories in real, school-based applications of educational technologies have been exceedingly rare (Cuban, 1986, 2001), the persistent belief that a new, more powerful technology such as the Internet will automatically change the face of education without concern for social, political, and pedagogical implications is difficult to dispel. New and more sophisticated technological devices are always being developed (such as mobile computing) and the rhetoric around their potential impact on education in popular media and even some reputable journals looms large. Abram's (2006) recent enthusiastic endorsement of iPods in education is typical:

I think that iPods and other more generic MP3 players are a bellwether technology.... To ignore iPods and their kin in the education space in 2006 is the same as ignoring the Web in 1996 or the Internet in 1986. You won't go extinct, but you won't evolve too quickly either.

If anything should have been learned from research in the field of educational technology by researchers and practitioners alike, it is that a tool itself will not change the educational system or even implicitly encourage new pedagogy. If the Internet and computers are going to reach their much-lauded potential as truly revolutionary tools, then something fundamental in the way educational technology research is done must change — and we believe this can occur. First, this change requires a shift in our concept of technology. Technology is much more than hardware. It is a process that involves the complex interactions of human, social, and cultural factors as well as the technical aspects. Second, it requires new directions in research goals, moving away from traditional predictive methods to long-term collaborations based on development goals.

## **Connecting education, research, and the technological condition**

We argue that educational researchers of all areas should be encouraged to move towards more systematic and collaborative methods of investigation that can promote research that makes a difference. In order to promote this agenda, two things must inform research in educational technology: first, an understanding of technology and technique as processes rather than artifacts; second, a resolute concern for the values, and principles guiding educational technology research. What Winner (1993/2003) says about social constructivists well describes most of the predictive research into educational technologies at the present:

... this perspective does not explore or in any way call into question the basic commitments and projects of modern technological society. The attitude of social constructivists seems to be that it is enough to provide

clearer, well-nuanced explanations of technological development...there is something very important missing here; namely, a general position on the social and technological patterns under study (p. 241).

Much research in educational technology still ignores the complex interaction between technological interventions, the roles of educational institutions such as schools and universities, the purposes of education, and the meaning of research. Many educational technology researchers adhere to a value-free discourse regarding the role of technology. There is a spotlight on the value of technology only to the extent that it has, or does not have an effect on learning-related variables. Indeed, it almost seems that many educational technologists have taken technological determinism as a given, and are simply trying to make the best of what is thrown at them by forces beyond their control. This positions educational technology researchers and practitioners at the end of the technological process, continuously testing new devices based on educational values that are not necessarily laudable.

If technology is recognized as a process rather than a mere artifact, then two things occur. First, researchers must begin to question their research methods due to the complexity of the environment under study. Investigations of how a “tool” does or does not affect educational outcomes are too simplistic. Second, researchers must question the values that are guiding research agendas, actively engaging with practitioners in constructing what constitutes valuable research in order to help direct technological development rather than react to it. We explore these two concerns in more detail below, followed by a discussion on how design-based research methods address these issues.

## Defining technology

Most educational technologists would accept the proposition that integrating technologies into an educational context is a complex task, partially because there are many stakeholders with differing respective values and interests. Fewer may be willing to concede that the Internet in itself is value laden. One could argue that computers and the Internet are inherently apolitical and value-free. After all, how could a computer promote any particular world-view? A device has no particular bias — it is up to humans to decide what purpose it should serve (for a discussion, see Pitt, 1987).

At this juncture, it becomes important to differentiate between the popular use of the term *technology*, and a more robust and accurate representation. The word is commonly used in the field of instructional and educational technologies to refer to electronic tools or devices such as the calculator, television, and the computer. This view of technology as a device prescribes educational technologists with a comfortable, albeit false, level of control and an easy, but ultimately inadequate, unit of analysis in their research pursuits.

This limited view of technology must be challenged at the definitional level. Technology is not a product and instead is a process: tools are merely a product of a technological system. A more inclusive definition of the term is offered by Hickman (2001), who uses Dewey’s pragmatism to describe technology as a process that involves the “invention, development, and cognitive deployment of tools and other artifacts, brought to bear on raw materials and intermediate stock parts, with a view to the resolution of perceived problems” (p. 26). While it might be broad in scope, it does well in describing the job that researchers and practitioners in educational technology regularly do: inquiry into techniques and tools in an effort to improve and refine the process of teaching and learning and, consequently, the design of learning environments.

The technological system is concerned with uncovering knowledge and information in so much as it leads to doing. These processes are planned, and the products that result from them are not the result of coincidence, though consequences might be unexpected. Technology can be seen as deterministic or as subservient to some other agent’s (human) control. While few would blindly ascribe to technological determinism, many naively assume the complete authority of man over the technological system (Ellul, 1980). One could create a parallel between this differentiation within educational technology research considering the distinction between types of “basic” and “applied” research, which differ based on the level of pragmatism involved in the research process (Hannafin, 2005; Reeves, 1995).

What is important here is to recognize that because of its pragmatic nature, technology cannot be considered to be value-free once it is recognized as both a process and a practice. As Ellul (1980/2003) contends, we cannot expect application to be judged as good or bad if we as researchers, from the onset, ignore the merits of moral judgment

within the research process. We maintain that educational technologists should not continue to simply investigate the impact or describe “best cases” in *post facto* applications of technological devices. This position makes them simply part of this technological system (Heidegger, 1977/2004), perpetually testing the appropriate uses of new technological devices in education.

More often than not, developments that occur outside of the educational arena are examined for educational affordances — in other words, attempting to examine the educational benefits of a new tool. There is nothing inherently wrong with the testing of new tools and techniques. Indeed, as noted earlier, that is much of what educational technology research does. Misguidance occurs as researchers get caught in a cycle of research without contemplating the merits of the investigation and the values implied by the tool or technique being used. Evidence of this trend can be seen in fifty-plus years of media comparison studies conducted by educational technology researchers to examine the influence of devices on educational achievement, with the most frequent result being “no significant differences” (Clark, 1983; Reeves, Herrington, & Oliver, 2004).

That educational technology is not value- or culture-free may be a hard proposition for some to accept. Neither education nor technology is neutral and unbiased (Freire, 1985; Hlynka, 2003). Their conception and application are guided by and provide guidance for political processes that are not necessarily grounded in principles that promote social good. Educational technologies are intricately connected with political agendas, economic gains, and social needs and consequences. Because of this, educational technologists should not be purveyors of “treatments” as if these devices and techniques were unbiased and value neutral. Computers and Internet access in schools are products of governmental policies that demand them (Department of Education, 1996, 2000), corporations that produce them, and numerous people who are often misinformed or ignorant about their purpose in education.

As part of the intricate socio-technical system that promotes the use of computers and other devices in education, educational technologists must begin to question and influence the *a priori* integration of these devices based on an investigation of its ends. As Borgmann (1984/2004) has discussed, the technological system upholds the division of ends and means. This is no different in education. Educational technologists are frequently more concerned with the possibilities of using a new technology (means), such as a newer course management system or the hottest wireless device, than seriously considering the ultimate aims of its use and its consequences.

As actors in the technological movement, educational researchers must take a critical stance towards technological development. Many of these concerns could be addressed by a serious media and technology literacy movement, but this has failed to develop into a cohesive effort in most schools (Amiel, 2006; ITEA, 2000; Petrina, 2000).

What is missing from the extant research in educational technology are questions of principle and value in regard to technological development. Hence, there is a need to add axiology (questions of quality or value) to the epistemology-theory-methodology-method thread that forms the basis of our educational inquiry (see Crotty, 1998). Design-based research calls for practitioners and researchers to engage in long-term collaborations (Reeves et al., 2004). It is necessary but not sufficient to connect research methods to compatible theoretical perspectives and epistemologies. Once a pedagogical stance is taken and we align it to a “way of knowing” (epistemology), then we must evaluate the *why* of what we want to know. The process of knowing in educational technology research is not disconnected from practice, and therefore implies change.

Far too often, researchers are tempted to adapt the educational environment to a new technique or device. For example, research using cell phones (music players, projection devices) in the classroom might have the noble objective of fostering better teaching and learning. But *better* is a value that must be disclosed. What does better teaching mean for the researcher and practitioner? To what end is this project being conducted? What are possible negative consequences? There should be clearer educational principles and foundations guiding the project, which then might demand a technological solution.

From the onset of any project, researchers into the field of educational technology must evaluate the principles that guide their research projects and the values that are promoted by their agendas. Researchers must not blindly accept the inherent values associated with technological development, and instead should seriously consider the nature of value in their practice (Koetting & Malisa, 2004). Gone unquestioned, the values promoted by the technological process are clear. Technology mediated by powerful interest groups is based on and promotes efficiency, speed, control, and reliability — values that primarily emphasize economic utility.

More often than not, the implementation of new technologies aims at making processes more efficient or more flexible, which is not necessarily valuable in its own right. Surely if research intends to effect change, we must realize that education is not simply about increasing the efficiency in the acquisition of knowledge and skills. As Postman (1995) has highlighted, “any education that is mainly about economic utility is far too limited to be useful, and in any case, so diminishes the world that it mocks one’s humanity” (p. 31). What values could be more exemplary of economic utility than the efficiency, speed, control, and reliability evidenced in the technological system? This conundrum indicates that researchers must find avenues to pursue valuable ends to the interventions they conduct in the name of education. We discuss the potential of design-based research as a framework to help us ask the right questions in educational technology research. This framework, which is based on meaningful practitioner-researcher connections from the onset of a research program, has the potential to address the issues highlighted above and guide the use of techniques and tools in education.

## Potential of design-based research

Design-based research (similar approaches have been termed design research, development research, and others) has recently received considerable attention by researchers in education as an emerging framework that can guide better educational research (Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006; Brown, 1992; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Design-Based Research Collective, 2003). Reeves (2006) outlines three cornerstone principles of this research framework:

“... addressing complex problems in real contexts in collaboration with practitioners; integrating known and hypothetical design principles with technological advances to render plausible solutions to these complex problems; and conducting rigorous and reflective inquiry to test and refine innovative learning environments as well as to define new design principles” (p. 58).

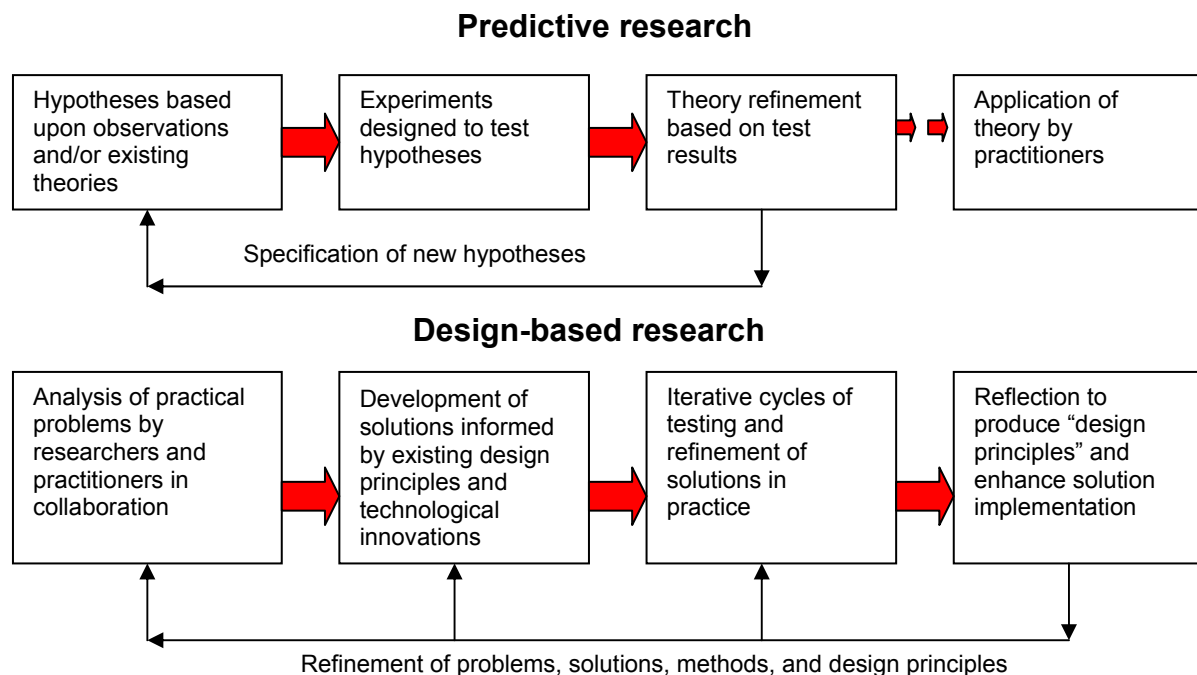


Figure 1. Predictive versus design-based research

The ultimate goal of design-based research to build a stronger connection between educational research and real-world problems. An emphasis is placed on an iterative research process that does not just evaluate an innovative product or intervention, but systematically attempts to refine the innovation while also producing design principles

that can guide similar research and development endeavors. This results in a cycle of research that is markedly different from what is currently pursued by many researchers in the field (Figure 1). In traditional empirical predictive research, a new technique or device is put to the test in a controlled environment. The time of engagement with the “stimulus” is usually limited because of time constraints. Iterations are encouraged in order to refine hypotheses, but commitment to iterative design is usually limited (one-shot studies).

Finally, researchers rarely engage directly with practitioners in the design process and, if this is done, participation is limited and occurs outside of the domain of practice. Practitioners are rarely part of the research design process, and are meant to reap the benefits of research when it is complete.

In contrast, we suggest that design-based research begin with the negotiation of research goals between practitioners and researchers (Figure 1). The practitioner is seen as a valuable partner in establishing research questions and identifying problems that merit investigation. Next, a design for the learning environment is proposed to address the concerns. This design could be a new set of strategies or it could be based on research gathered from previously tested design principles. The design-based researcher is humble in approaching research by recognizing the complexity of interactions that occur in real-world environments and the contextual limitations of proposed designs. The development of design principles will undergo a series of testing and refinement cycles. Data is collected systematically in order to re-define the problems, possible solutions, and the principles that might best address them. As data is re-examined and reflected upon, new designs are created and implemented, producing a continuous cycle of design-reflection-design. The outcomes of design-based research are a set of design principles or guidelines derived empirically and richly described, which can be implemented by others interested in studying similar settings and concerns. While the ultimate objective is the development of theory, this might only occur after long-term engagement and multiple design investigations.

While the methods used to conduct design-based research are not new, the intentions and lifecycle proposed by its framework are certainly in a unique position to address the complexities inherent in educational technology research. Critiques and analyses of design-based research have been postulated by others and are beyond the scope of this paper. Our focus is on the contribution of design-based research in studying the complexity of technology as a process and shaping the question of value of research by establishing relationships between practitioners and researchers.

How does a new research framework such as design-based research address the conception of technology as a process and the issue of value in educational technology research? Five characteristics of design-based research identified by Van den Akker et al. (2006) will be examined to address this question. They are as follows: interventionist, iterative, process oriented, utility oriented, and theory oriented.

The understanding of technology as a process greatly increases the complexity of the integration of tools into educational environments. Educational technologies become more than simply an independent variable in a study of student learning. Integrating technologies into the classroom leads to substantial changes in social organization, student-teacher relationships, and a myriad of other factors that cannot be investigated successfully by predictive research. Researchers must make a commitment to conducting interventionist research in real-world contexts such as schools, accepting the complexity of the setting. As Kafai (2005) contends, schools can become “living laboratories” in which researchers investigate in real-world settings while attempting to control for critical variables identified through theory and previous research.

Still, the sheer number of variables is indeed so many that one-shot studies of impact would lead to very limited insight. Design-based research calls for iterative cycles of study that lead to a better understanding of the process of intervention (process oriented). Indeed it would be idealistic to expect significant and transferable results from a one-time study of a technological intervention. Using iterative cycles of design and re-design allows for the investigation of these critical variables and limitations, generating more transferable and useful results.

Nelson, Ketelhut, Clarke, Bowman, and Dede (2005) describe this process in the design of a multi-user virtual environment called River City, used to teach children about water pollution. They tested multiple iterations of design with a variety of different groups of children while making systematic changes to certain aspects of design and documenting its effects. Multiple iterations allowed for design changes and testing with a variety of groups, lending greater transferability to the design itself. If enough designs are implemented and evaluated systematically, the



designs themselves have the potential to contribute to, or generate theory or conjectures (for a critique of the issue see Kelly, 2004). Barab, Arici, and Jackson (2005) describe the development of learning engagement theory as part of the design of Quest Atlantis, an immersive online environment. Systematic investigation of design factors over a sustained period of time allowed for a theory of engagement to emerge. This theory has particular value since it emerged from design within the ecology of school, as opposed to being generated in more sterile or controlled conditions.

Finally, the strong commitment to intervention in real-world settings and its orientation towards utility bring forward the concern with values in educational technology research. As a principle, design-based research advocates for practitioner-researcher partnerships. If taken seriously, such partnerships have strong implications for the position and power of educational technology researchers.

Technology can be seen as an autonomous phenomenon that perpetuates and promotes itself, subordinating political decisions to an on-going cycle in search of better and more efficient ways. Feenberg (2002, 2003) rejects the perspective that technology is inherently autonomous, suggesting instead that the technological system is historically serving the needs of a particular hegemony. In order to break with the hegemony, Ellul (1992) calls for a revisiting of the type of democratic institution that renews the power of individuals in guiding their public and private lives (see also Hickman, 2001). Here, Ellul (1992) makes a clear and strong connection between education and democracy in a technological society, arguing that the public “must be given information that allows for free decisions, not ones based solely upon a menu of options served up by technicians” (p. 44). Hickman (2001) expands on this view, arguing that a Deweyan take on technology would promote the education of citizens to encourage their involvement in the deliberation of the design and implementation of technological tools as well as the research that informs their application. This process would lead to a more balanced collaborative deliberation between experts and non-experts inquiring into societal tribulations. In order to make and interpret these decisions, dialogue and discussion are necessary.

Let us bring this debate into the realm of educational technology research, seeing school as a microcosm of society. Debate and dialogue in decision-making are fundamental to a democratic society and democratic practice (Parker, 1996), why should this be any different in the realm of research? This brings forward the necessity of recognizing the voice of practitioners as invaluable to the design process. Researchers should not see themselves as external technocrats, bringing solutions to envisioned school problems. The issues addressed by educational researchers in school must emerge from the school itself through its constituents. These problems must be negotiated between school members and researchers. It is important to highlight that this is not a reversal of directionality. As Dede (2005) points out, there is an important balance as to what practitioners consider to be pressing issues, and what researchers (and research) have identified as problems. This conflict of values and ideas opens up a valuable space for debate. Teachers become active partners in identifying priorities for research and contributors throughout the research process itself. Proponents of action research have long recognized the importance of the teacher/practitioner in research. Design-based research brings this framework into its fold, and adds the possibility of not only solving the practice-oriented problems addressed by action research, but also identifying reusable design principles.

The introduction of cooperation between researchers and practitioners at an early stage of research is a unique approach to improving both the value of educational technology research and its potential to direct technological development in schools. The reality check of engaging directly with practitioners and school environments has the potential to eliminate much research that is not valuable or socially responsible. Design-based research does not in itself demand a particular agenda for research. While researchers or teachers could ask irrelevant questions, having a serious negotiation and debate over the research agenda greatly increases the possibility that the right questions will be asked — questions that will lead to research that produces useful and applicable knowledge addressing the needs of teachers. The concerns of practitioners, if given full consideration, rarely address irrelevant issues. The job of researchers and practitioners is to cooperatively negotiate what is worthy to investigate. This negotiation in turn can help eliminate the type of studies that investigate the “affordances” of devices for the sake of novelty. The attempt to begin research on a new tool can be critically assessed through the constraints of a real-world environment and the voice of the practitioners who can help evaluate the usefulness of such tools. Cycles of design informed by real-world scenarios can help clearly identify which interventions merit adoption and in which contexts this should occur. Considering that research should help decide on adoption patterns for educational technology in schools, researchers would have an active role within the technological system, helping direct the development and acceptance of new tools and techniques.

For example, Amiel, McClendon, and Orey (2007) describe a four-year program in which researchers engaged directly with schoolteachers in Brazil and the United States. Teachers, pre-service teachers, and researchers collaborated in the design of learning environments. Researchers were interested in promoting cross-national dialogue and discussion on all areas of the curriculum as part of a democratic education framework by connecting public schools in both countries through internet-based technologies. Within this broad research and practice agenda, the concerns of schoolteachers were starting points for month-long discussions on how and why these connections would take place. The “give and take” of researchers and practitioners set the scope for what would be implemented. The dialogue and discussion between stakeholders, a cornerstone of democratic practice, is too often missing from the agenda in educational technology research. As a result of this study, Amiel, McClendon, and Orey (2007) present a model of collaboration that can occur in projects involving university staff, students, and public schools across national borders. This model evolved over time and was the product of multiple project iterations. It is by no means a theory of collaboration but a conjecture (see Kelly, 2004) or design principle that will evolve and mature over time.

## Conclusion

The aforementioned call for design-based research presents educational researchers with a conundrum: if we persist in believing in education and technology as value-free, we should not attempt to engage in design-based research and should instead resign ourselves to perpetuating research that effects no systematic change. We may hide our lack of concern for impact behind the veil of academic freedom. But if the case for the new design-based methodologies is sound, then research and practice can become intertwined, and as a result, it becomes impractical and indeed *ungrounded* to promote the kinds of impartial, unengaged research that dominates the published literature.

Researchers in the field of educational technology can begin to look away from the short-term objectives of their individual projects. In order to escape the anti-humanistic values often promoted by technological development, educational technologists must recognize the transformational potential of their profession. A primary responsibility of researchers in the field should be to limit their investigation of means and contemplate educational ends or aims, making them explicit in the process of an investigation.

Design-based research provides a cycle that promotes the reflective and long-term foundation upon which such research can be undertaken. Educational technology researchers should be concerned with examining the technological process as it unfolds in schools and universities and its relationship to larger society. By carefully considering their ends and selecting an appropriate methodology, researchers in our field will be better prepared to determine their values, make their agendas explicit, and promote democratic practice.

This pursuit of socially responsible research may be more important than ever. We live in an age when a U.S. president issues his first veto in five years to ban funding for embryonic stem cell research, surrounded by children born from “rescued” embryos, ignoring the fact that none of the children would have been born without the contributions of earlier generations of embryonic researchers. We live in a world of melting glaciers and rising seas, when more people appear to believe in angels and ghosts than in global warming. Is it too simplistic to suggest that educational technology researchers might have a role in combating such global ignorance? Perhaps so. But we think not.

## Acknowledgements

The authors would like to thank Mike Orey, Xornam Apedoe, and Jo McClendon for comments on previous versions of this article.

## References

- Abram, S. (2006). The proof is in the podding. *Multimedia & Internet @ Schools*, 13(3).
- Amiel, T. (2006). Mistaking computers for technology: Technology literacy and the digital divide. *AACE Journal*, 14(3).

- Amiel, T., McClendon, J., & Orey, M. (2007). A model for international collaborative development work in schools. *Educational Media International*, 44(2), 167–179.
- Barab, S., Arici, A., & Jackson, C. (2005, January/February). Eat your vegetables and your homework: A design-based investigation of enjoyment and meaning in learning. *Educational Technology*, 15–21.
- Bennett, F. (1996). Computers as tutors. Retrieved September 10, 2006, from <http://www.firstmonday.dk/issues/issue6/intro/index.html>
- Borgmann, A. (1984/2004). Focal things and practices. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 115–136). New York: Rowman & Littlefield.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions. *Journal of the Learning Sciences*, 2(2), 141–178.
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445–459.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9–13.
- Crotty, M. J. (1998). *The foundations of social research: Meaning and perspective in the research process*. Thousand Oaks, CA: Sage.
- Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York: Teachers College Press.
- Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Cambridge, Massachusetts: Harvard University Press.
- Dede, C. (2005). Commentary: The growing utilization of design-based research. *Contemporary Issues in Technology and Teacher Education*, 5(3/4), 345–348.
- Department for Education and Skills. (2003). Survey of information and communication technologies in schools. Retrieved September 10, 2006, from <http://www.dfes.gov.uk/rsgateway/DB/SBU/b000421/index.shtml>
- Department of Education. (1996). Getting America's students ready for the 21st century. Retrieved September 8, 2006 from <http://www.netc.org/cdrom/tlc/html/execsum.html>
- Department of Education. (2000). E-learning: Putting a world-class education at the fingertips of all children. Retrieved November 10, 2003, from <http://www.ed.gov/about/offices/list/os/technology/reports/e-learning.html>
- Department of Education. (2005). FY 2006 budget summary. Retrieved September 10, 2006, from <http://www.ed.gov/about/overview/budget/budget06/summary/index.html>
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8.
- Dickard, N. (Ed.). (2003). *The sustainability challenge: Taking edtech research to the next level*. Benton Foundation.
- Ellul, J. (1980). *The technological system*. New York: Continuum.
- Ellul, J. (1980/2003). The “autonomy” of the technological phenomenon. In R. C. Scharff & V. Dusek (Eds.), *Philosophy of technology: The technological condition* (pp. 386–397). Malden, MA: Blackwell Publishing.
- Ellul, J. (1992). Technology and democracy (D. Lovekin & M. Johnson, Trans.). In L. Winner (Ed.), *Democracy in a Technological Society* (pp. 35–50). Dordrecht, The Netherlands: Kluwer.
- Feenberg, A. (2002). *Transforming technology*. New York: Oxford University Press.
- Feenberg, A. (2003). Democratic rationalization: Technology, power, and freedom. In R. C. Scharff & V. Dusek (Eds.), *Philosophy of technology: The technological condition* (pp. 652–665). Malden, MA: Blackwell.
- Freire, P. (1985). *The politics of education*. South Hadley, Massachusetts: Bergin & Garvey.
- Gartner Inc. (2004). Hype cycle special report. Retrieved February 5, 2004, from <http://www4.gartner.com/pages/story.php.id.8789.s.8.jsp>.

- Hannafin, M. (2005). Disciplined inquiry and research in computer-supported learning. *Computers in Human Behavior*, 22, 149–153.
- Heidegger, M. (1977/2004). Question concerning technology. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 35–52). New York: Rowman & Littlefield.
- Hersh, R. H., & Merrow, J. (Eds.). (2005). *Declining by degrees: Higher education at risk*. New York: Palgrave Macmillan.
- Hickman, L. A. (2001). *Philosophical tools for technological culture*. Bloomington, IN: Indiana University Press.
- Hlynka, D. (2003). The cultural discourses of educational technology: A Canadian perspective. *Educational Technology*, 43(4), 41–45.
- ITEA. (2000). Standards for technology literacy [executive summary]. Retrieved September 10, 2006, from <http://www.iteaconnect.org/>
- Jeffreys, M. V. C. (1955). *Beyond neutrality: Five essays on the purpose of education*. London: Pitman Press.
- Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational technology research and development*, 39(3), 5–14.
- Jonassen, D. H. (2003). The vain quest for a unified theory of learning. *Educational Technology*, 43(4), 5–8.
- Jones, G. R. (1996). *Cyberschools: An education renaissance*. Englewood, CO: Jones Digital Century.
- Kafai, Y. B. (2005, January/February). The classroom as “living laboratory”: Design-based research for understanding, comparing, and evaluating learning science through design. *Educational Technology*, 28–34.
- Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *The Journal of the Learning Sciences*, 13(1), 115–128.
- Koetting, J. R., & Malisa, M. (2004). Philosophy, research, and education. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (2nd ed., pp. 1009–1020). New York: MacMillan.
- Light, J. S. (2001). Rethinking the digital divide. *Harvard Educational Review*, 71(4), 709–733.
- Loveless, T., & Diperna, P. (2000). *The Brown Center report on American education: How well are American students learning? Focus on math achievement*. Washington, DC: Brookings Institutions.
- Nelson, B., Ketelhut, D. J., Clarke, J., Bowman, C., & Dede, C. (2005, January/February). Design-based research strategies for developing a scientific inquiry curriculum in a multi-user virtual environment. *Educational Technology*, 21–28.
- Office of Technology Assessment. (1995). *Teachers and technology: Making the connections*. Retrieved April 23, 2003, from <http://www.wws.princeton.edu/cgi-bin/byteserv.prl/~ota/disk1/1995/9541/9541.PDF>
- Oppenheimer, T. (1997). The computer delusion. *Atlantic Monthly*, 280, 45–62.
- Oppenheimer, T. (2003). *The flickering mind*. New York: Random House.
- Parker, W. C. (1996). *Schools as laboratories of democracy*. Albany, NY: SUNY Press.
- Perelman, L. J. (1992). *School's out: Hyperlearning, the new technology, and the end of education*. New York: William Morrow.
- Petrina, S. (2000). The politics of technological literacy. *International Journal of Technology and Design Education*, 10, 181–206.
- Pitt, J. C. (1987). The autonomy of technology. In P. T. Durbin (Ed.), *Technology and Responsibility* (pp. 99–114). Boston: D. Reidel Publishing Company.
- Postman, N. (1995). *The end of education* (1st ed.). New York: Alfred A. Knopf.
- Postman, N. (2003). Questioning media. In M. S. Pittinsky (Ed.), *The wired tower: Perspectives on the impact of the internet on higher education* (pp. 181–200). Upper Saddle River, NJ: Prentice Hall.

- Reeves, T. C. (1995). *Questioning the questions of instructional technology research*. Paper presented at the Annual Conference of the Association for Educational Communications and Technology, Anaheim, CA.
- Reeves, T. (2006). Design research from a technology perspective. In J. V. D. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 52–66). New York: Routledge.
- Reeves, T. C., Herrington, J., & Oliver, R. (2004). A development research agenda for online collaborative learning. *Educational Technology Research and Development*, 52(4), 53–63.
- Rescher, N. (1980). *Unpopular essays on technological progress*. Pittsburgh: University of Pittsburgh Press.
- Saettler, P. (1990). *The evolution of American educational technology*. Englewood, CO: Libraries Unlimited.
- Salomon, G. (2002). Technology and pedagogy: Why don't we see the promised revolution? *Educational Technology*, 42(2), 71–75.
- Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Introducing education design research. In J. V. D. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 3–7). New York: Routledge.
- Wang, F., & Reeves, T. C. (2003). Why do teachers need to use technology in their classrooms? Issues, problems, and solutions. *Computers in the Schools*, 20(4), 49–65.
- Winner, L. (1993/2003). Social constructivism: Upon opening the black box and finding it empty. In R. C. Sharff & V. Dusek (Eds.), *Philosophy of technology: The technological condition* (pp. 233–243). Malden, MA: Blackwell.

## Randomised Items in Computer-based Tests: Russian Roulette in Assessment?

**Anthony M. Marks**

Faculty of Engineering, Information Technology & the Built Environment, Summerstrand North Campus  
Nelson Mandela Metropolitan University // [anthony.marks@nmmu.ac.za](mailto:anthony.marks@nmmu.ac.za)

**Johannes C. Cronje**

Faculty of Informatics and Design, Cape Peninsula University of Technology // [johannes.cronje@gmail.com](mailto:johannes.cronje@gmail.com)

### ABSTRACT

Computer-based assessments are becoming more commonplace, perhaps as a necessity for faculty to cope with large class sizes. These tests often occur in large computer testing venues in which test security may be compromised. In an attempt to limit the likelihood of cheating in such venues, randomised presentation of items is automatically programmed into testing software, such that neighbouring screens present different items to the test-taker. This article argues that randomisation of test items can be a disadvantage to students who were randomly presented with difficult items first. Such disadvantage would violate the American Psychological Association's published guidelines concerning testing and assessment that call for the principle of fairness for test-takers across diverse test modes. Owing to the smallness of the chance of a student being randomly assigned difficult items first, it may be hard to prove such disadvantage. However, even if only one test-taker is affected once during a high-stakes test, the principle of fairness is compromised. This article reports on four instances out of about 400 in which students may either have been unfairly advantaged or disadvantaged by being given a series of easy or difficult items at the beginning of the test. Although the results are not statistically significant, we conclude that more research needs to be done before one can ignore what we have named the Item Randomisation Effect.

### Keywords

Computer-based tests, Fairness, Cheating, Randomisation of items, Anxiety

### Introduction

An important security feature of computer-based multiple-choice testing is that test items are randomised to prevent students working at adjacent computers from copying. The downside of such randomisation, however, is that it prevents planned sequencing of items, which is commonplace in the paper-based equivalent. A test constructor may, for instance, place easier items at the beginning of the test to build confidence in the test-taker, and place the most difficult items at the end so that slower students' time is not wasted by attempting items that are beyond their ability. In other cases, a student may choose to go through the test first and select the easier items, leaving the difficult ones for last. Randomising items does not accommodate a test user or a constructor who wishes to ensure that items progressively become tougher. Although navigation through computer-based tests is possible, it is certainly an inconvenience to test-takers who want to leave tougher items for last. In this study we wanted to know if it would be possible to identify students who were disadvantaged in a test because they had been randomly assigned the difficult items first.

The purpose of this study is to determine the effect of current computer-based testing practice on student performance, particularly with respect to the random item sequencing algorithm. The purpose of this algorithm or software code is to ensure that each test-taker is administered the test items in a sequence different from that of any other test-taker. This effect will not affect all test-takers equally and is like a form of Russian roulette, a dangerous game of chance, played in high-stakes testing contexts. Russian roulette does not cause injury to all the participants, yet is deadly to the one who pulls the trigger last. One test-taker affected by this form of unfairness in assessment is one too many. Test security may be enhanced, but is test fairness for some test-takers compromised?

Paper-and-pencil tests still account for a major portion of any student's final result. This is true from the early school years to the final years of postgraduate studies, during which time one can safely estimate that at least ten years of school achievement was assessed with the aid of written tests and examinations. Good test-taking habits will ensure that candidates perform according to their level of preparedness (Glenn, 2004). The test-taking skill of particular interest and relevance to this paper is the one that requires candidates to select their perceived easy questions first

and answer these before attempting their perceived tougher questions (“Use Parent Nights,” 2004). In paper-based tests, this is easily achieved by simply paging back and forth through the test.

Now with the advances in personal computer technology and huge investments in evaluation and testing software (Billings, 2004; Harding, 2001; Varughese, 2005), computer-based testing is becoming commonplace. Candidates who have good test-taking skills with regard to paper-based tests will still outperform candidates without these skills in computer-based tests. Or will they? Algorithms that randomise the order in which the test items are presented to each candidate automatically control certain computer-based test assessments.

Although randomisation may reduce the security risk (Pain & Le Heron, 2003) of adjacent students’ copying from or aiding one another, it may unfairly increase test anxiety for some of the candidates. Navigation through numerous test items in a computer-based test is not conveniently achieved by candidates, so randomly receiving several difficult items consecutively may unduly stress a candidate. Increased anxiety at any stage during the test for whatever reason is likely to have a negative effect on that person’s performance for the remainder of the test (Lufi, Okasha, & Cohen, 2004; Supon, 2004). Obviously, the sooner the sequence of difficult items is presented the more pronounced the effect it may have upon the remainder of the test. The main question of this study is “Can instances be found where randomisation of items in a computer-based test unfairly disadvantaged any of the test-takers in any way?”

The following sub-questions drove the study:

- What constitutes “normal performance” for each student in the sample?
- Did any candidates’ performance differ significantly from their “normal performance”?
- Were these candidates presented with consecutive randomly sequenced difficult items?
- Can this deviation be attributed to the randomisation of items presented to the student?

For the purpose of this study, the authors felt that primacy and recency effects on memory and cognition were not directly relevant. Primacy is a term from cognitive psychology that is used to explain the increased likelihood for accurate recall of the initial items of a list of items. Recency is a term used to explain the increased likelihood for people to accurately recall the items occurring at the end of the list. Studies of these effects (Bemelmans, Wolters, Zwinderman, ten Berge, & Goekoop, 2002; Talmi & Goshen-Gottstein, 2006) are more concerned with understanding the workings of memory and recall and offset discussion around long-term and short-term memory processes. However, in this study, a set of test items comprise different multiple choice questions, not a list of words and or numbers to be recalled in the correct sequence. The items or questions in this study are randomised so that test-takers are administered the same test, but the randomised sequence of items may subtly cause some candidates to be disadvantaged in comparison to the rest of the candidates taking the test.

The focus of this study is on the potential threat to fairness that randomisation of test items may cause to any one test-taker. This is especially important as item randomisation is done in high-stakes testing scenarios, for example, entrance examinations to tertiary institutions, promotion or retention of children in schools, or selection of potential candidates for a particular job vacancy (Russell, Goldberg, & O’Connor, 2003; Zenisky & Sireci, 2002). Consider what it would be like to be the test-taker and have the items presented in a manner that causes you to perform poorly in such an assessment, and to be disqualified from further consideration for the institution/post for the wrong reason. Hence, even if only one test-taker is unfairly administered a test, it is one too many.

## Literature survey

Tobias (1985) hypothesized that test anxiety reduced the cognitive ability required to solve problems, thus leading to poor results. On the other hand, a student with good test-taking skills needed less cognitive capacity to spend on the physical elements of the test and could therefore concentrate more on recall of the actual learning content. Bierenbaum (2007) built upon Tobias’s work and identified a perceived alignment between instruction and assessment. Beirenbaum argued that students come to expect a certain style of test and answer in a certain way. Deviations from such expectation lead to test anxiety.

This study seeks to show that a form of test anxiety may exist in computer-based tests due to randomisation of the items. This randomisation may affect a small percentage of test-takers in a sample and therefore may be seen as unfairly disadvantaging those candidates.

Fairness for everyone taking tests is the underpinning principle of the various publications on testing and assessment available from the American Psychological Association (APA, 2004; Turner, DeMers, Fox, & Reed, 2001). These guides are specifically used to ensure standardised tests and interpretations of test-taker abilities made from such tests are accurate and fair (Turner et al., 2001). Test developers and users are defined in the guidelines, and they are regarded as the stakeholders specifically tasked with ensuring that the guidelines are followed, and that fairness for all test-takers is achieved. Another publication has evolved from the guides to inform all three stakeholders, viz. developers, users, and test-takers, of the rights and responsibilities of test-takers (APA, 1988).

From the literature, fairness is seen to be a fundamental principle. Randomised sequencing prevents cheating, and test-taking skills involve a number of activities, some of which are compromised by randomised sequencing. Test stress negatively impacts test results, while navigational control enhances test performance. No literature could be found that comments specifically upon the sequencing of computer-based test items. However, Sternberg (1998) stresses the importance of metacognition as a part of what makes an expert student. We could argue that randomised sequencing impairs metacognition because it distracts from the holistic nature of a test.

Pain and Le Heron (2003) reported that randomised sequencing of test items was successful in preventing cheating in computer-based tests. In one of the scenarios, they even allowed the computer-administered test to randomly select different items from the question databank such that each student had a different collection of items presented to them for the assessment. Not surprisingly, this ensured that no students could cheat, but as the authors reported, the concern about the fairness of such a solution is questionable and they returned to allowing random sequencing of a set of test questions, such that all test-takers essentially took the same test. However, Pain and Le Heron did not consider the harmful effects of unfairly presenting sequences of difficult items early in the test.

Glenn (2004, p. 62) advises test takers to “answer the easiest questions first. Completing the sure-thing questions first boosts student confidence from the outset.” The literature does not say what happens to student confidence or anxiety if this important test-taking skill is ignored. We would like to deduce that the opposite effect will occur, that student confidence wanes and test anxiety increases. In computer-based testing contexts, several consecutive difficult items presented to a test-taker will increase that test-taker’s anxiety level. Therefore, that candidate is unfairly administered (Turner et al., 2001) the test in comparison with all the other candidates that were fortunate not to be randomly presented with such a sequence of difficult test items. This is clearly in contravention of the APA requirements that all tests are to be administered in a standardised manner such that all test-takers are given an equal opportunity to provide evidence of their abilities in that test (2004).

The above assumes that a sequence of difficult items will indeed cause increased anxiety. It also assumes that increased anxiety has a negative effect on student performance for the remainder of the test (Black, 2005). Cognitive ability decreases during states of tension and increased anxiety (“Reduce Test Anxiety,” 2005; Dutke & Stöber, 2001; Hancock, 2001). A review of literature pertaining to computer-based tests found various studies relevant to computer anxiety (Lufi et al., 2004; Supon, 2004; Tseng, Tiplady, Macleod, & Wright, 1998) and how it adversely affected student performance. Computer anxiety is prevalent in persons who seldom use computers in their daily lives, hence their nervousness when using computers. This computer anxiety compounds the natural anxiety caused by the need to perform adequately in a test setting, as stated by Bugbee Jr. (1996, Specific Research Studies section, para. 17): “Anxiety is quite real and can gravely affect a test taker. It must be dealt with.”

Various studies have already managed to identify diverse factors that cause significant differences in student performance to be observed across modalities, specifically paper-based versus computer-based test modes (Bugbee Jr., 1996; Carlson & Smith Harvey, 2004; Hoff, 1999). These differences are solved directly. For example, the ability to return to any item and edit its answer has helped ensure some equivalence across test modes (Ferguson, Kreiter, Peterson, Rowat, & Elliott, 2002). If a solution is not easily implemented, then the APA guidelines allow for scores to be adjusted such that the adjusted scores are fair representations of the test-takers when compared with persons taking tests in other test modes (Russell et al., 2003). However, the item randomisation test mode effect identified by this study has not yet been studied.



The paper and pencil mode conveniently allows test-takers to read through the entire set of items before choosing to attempt the easier ones first, as prescribed by those advocating this approach as a good test-taking tactic (“Use Parent Nights,” 2004; Glenn, 2004; Staber & Pekrun, 2004). However, randomised sequencing of test items in computer-based test assessments is not determined by the test-taker. Navigation between items is not as convenient for computer-based test assessments as it is for paper and pencil, and this forms the basis of the argument that students in computer-based tests are all at a relative disadvantage. This navigation mode effect has been studied (Ferguson et al., 2002), and as all test-takers in this modality are equally affected, it was readily shown that by allowing candidates more control, the disadvantage could be somewhat negated. However, it then became obvious that students then needed extra time to navigate back and forth through the test items as compared to students sitting the same test in a paper-based mode.

The mode effect described in our study will only affect a small number of test-takers randomly in a computer-based test, not all of the candidates. It follows that this effect is not likely to be easily noticed. Hence the need for this study, as the available literature does not include any studies in this regard.

## **Method**

### **Overview**

In this study we investigated the performance of 103 third-year students of veterinary science in four tests out of five that were presented during a year-long course. These tests were completed in the year prior to the research commencement as it was hoped that individual occurrences of the randomisation effect would be found from existing data. From an ethical perspective it is important to note that no students were disadvantaged as a direct result of this investigation. This research is an ex-post facto study done on existing computer-based test data. Further, the research is a pilot study done to investigate the potential or need for rigorous experimental research studies to be designed to properly investigate what has now been called the Item Randomisation Effect. No students were interviewed before, during, or after the collection of the data as this was outside of the intended scope of the pilot study. Similarly none of the tests was manipulated in any way. The data was simply collected from the computer-based testing department at least one year after the tests were administered as this was believed to be sufficient for the purposes of this pilot study. The data were analysed and the normal performance of each student was established as described below. The test data were processed until the relevant variables were ready for analysis. Next, the students who deviated significantly from their normal performance were flagged, and their test experience was then analysed in detail. The graphical representation of selected students is included later in the paper. It must be stressed that this research was ideographic rather than nomothetic — we were looking for specific instances rather than considering the over-all test performances.

### **Normal performance for this study**

The performance of a candidate who was presented with difficult items during the initial stages of any of the tests needs to be compared with the normal performance of that candidate in similar assessments, and significant deviations in the performance of that candidate need to be reported. One obvious record of normal performance could be obtained from each student’s academic record and the average obtained thus far in his or her academic career. However, this was not satisfactory because deviations above and below the average are normal, based upon the individual student’s aptitude and motivation levels for different courses. We decided, rather, to find a sample of students who had sat more than two computer-based test assessments in one year-long course. The average obtained for each candidate for the particular course would, for the purpose of this study, be considered the normal performance of the student.

### **Defining difficult items for this study**

The Difficulty Index and the Discrimination Index are two of the indices that are readily available in computer-based tests. Most testing software calculate these as standard features to assist users in designing and retaining quality test items, yet flag problematic items that should either be edited or discarded (Alessi & Trollip, 2001; Reise & Henson,

2003). In the context of this study, the difficulty index is the obvious data that allows one to check the correlation between the sequence of items and the relative difficulty of those items. The discrimination index is useful but of lower significance for the purposes of this pilot study, and is therefore not considered any further in this paper. The difficulty index is calculated by dividing the number of correct responses for an item by the total number of attempts made to answer the item (Reise & Henson, 2003). The index can range between zero and unity. A zero index value means that none of the test-takers could correctly answer the item. A difficulty index of unity means that all the test-takers attempting the item chose the correct option (or key). Assuming the items were not compromised in any way, a zero index implies a difficult item and, conversely, a unity index implies an easy item. Once used in a test, each item can now be ranked according to its difficulty index and thus labelled and stored in the question databank.

### **Sample data selection for this study**

The computer-based testing section of the University of Pretoria assisted us in finding a set of data that satisfied the requirements of numerous tests in one subject in a particular year for as large a sample of students as possible. The chosen sample consisted of four sets of test results for 103 veterinary students in their third year of study in 2004. One of their modules required them to sit four computer-based test assessments. Questionmark 3.2 was the software used to administer the tests. The data available for each test consisted of four text files generated by Questionmark after the completion of each test. In addition, we were supplied a portable document format (pdf) copy of each set of test items. Permission to use the data was granted by the Veterinary Sciences Faculty. Considerable effort was required to organise the four text files such that the data required for this study could be analysed. This is because the required variables appeared in different text files.

### **Data cleaning**

Four sets of existing test data from a sample of students for one course in their academic year was used for this study. They are labelled Test 1, Test 2, Test 3, and Test 5. Test 4 data was eliminated because problems during the test led to its being postponed until the next day. The test items were all changed as many of the candidates had seen some of the items, so a new test, Test 5, was created. Thus only four sets of test data were useful.

The normal performance for each test-taker was taken as the average scored for the four tests in this course, as was explained earlier. Next, it was important to obtain the sequence of items as randomly presented to each test-taker. The last step required the difficulty index of each item to be recorded adjacent to each item for the candidate. The student score for each item was also included as was the discrimination index and other potentially useful information. The process was repeated for each of the four sets of test data. The data was then ready for analysis.

### **Discussion from findings**

Detailed analysis of the particular students who had test scores varying by at least 15 percent does seem to show correlations between the sequences of item difficulty and the effect on students' performance on the test. The literature indicates that beginning with easy items in a test is one habit likely to have a positive effect on student grades (Glenn, 2004). In this study, students who scored significantly higher than usual in Test 1 tended to be presented with the easy items early in the test. Conversely, those presented the difficult items early in the four tests tended to perform significantly below their normal levels.

The figures that follow, one from each test, are the most striking examples of the trends that become apparent when one studies the data pertaining to the candidates who showed a significant difference in performance from their normal performance for this module across the four tests. The following conventions are followed in the figures:

- (1) Diamonds indicate the difficulty index for each item
- (2) The horizontal line labelled "average difficulty" is also the class average scored for the particular test
- (3) Diamonds 10% above the line of average difficulty are considered "easy"
- (4) Diamonds 10% below the line of average difficulty are considered "difficult"
- (5) Diamonds within 10% of the line of average difficulty are considered "average"
- (6) The candidate's deviation from normal is indicated under the chart's title.

Figure 1 shows an example of a candidate who scored 24 percent above his/her normal. Closer inspection of the chart shows that for the first eight items, six were easy, one difficult, and one of average difficulty. This candidate was presented the items in an almost ideal sequence, easy to difficult, and scored significantly above his/her normal, which supports the literature pertaining to attempting easy items first (Glenn, 2004; Supon, 2004).

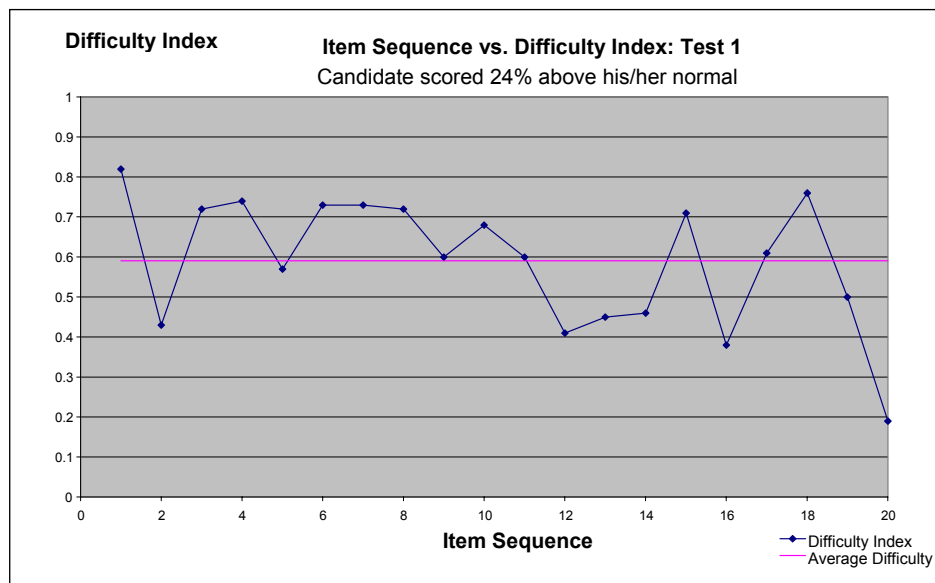


Figure 1. Example from Test 1

Figure 2 shows an example of a candidate who scored 19 percent below his/her normal level. For the first eight items, two were easy, two difficult, and four of average difficulty. The ideal would have been easy items in the beginning, average items in the middle, and difficult items at the end. It is important to remember that the assumption that the difficulty index indicates the degree of difficulty is valid for the class group as a whole but not necessarily true for each individual student. This particular student got the second and fifth items correct, inferring that the other six items of the first eight were perceived to be difficult for this particular student, hence the potential for increased anxiety as described throughout this paper.

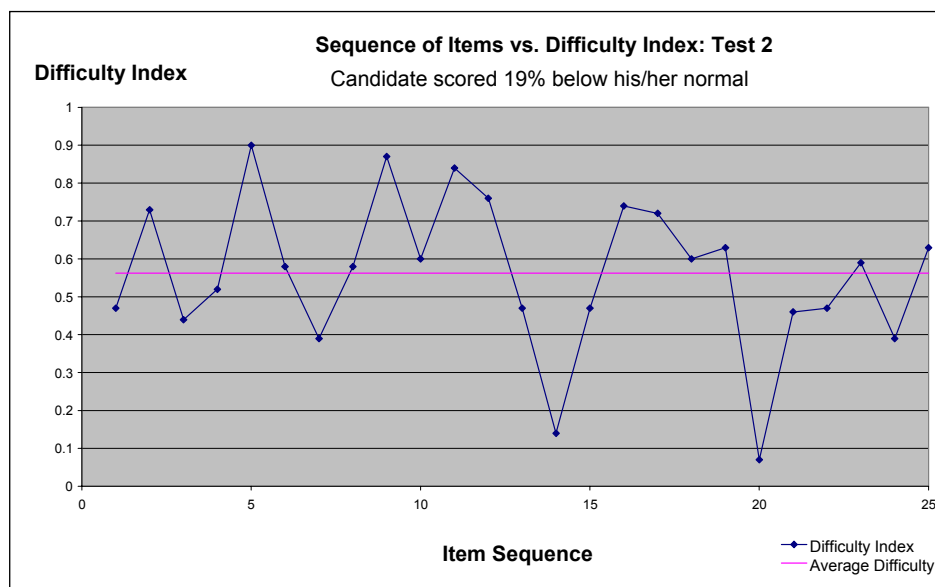


Figure 2. Example from Test 2

In Figure 3 the candidate scored 15 percent below his/her normal level. In first third of the test (the first 16 items), four questions were easy, six difficult, and six of average difficulty. Most of these questions ranged from average difficulty to difficult, which is ideal for the middle third of the test. Did this random sequence of test items cause increased anxiety?

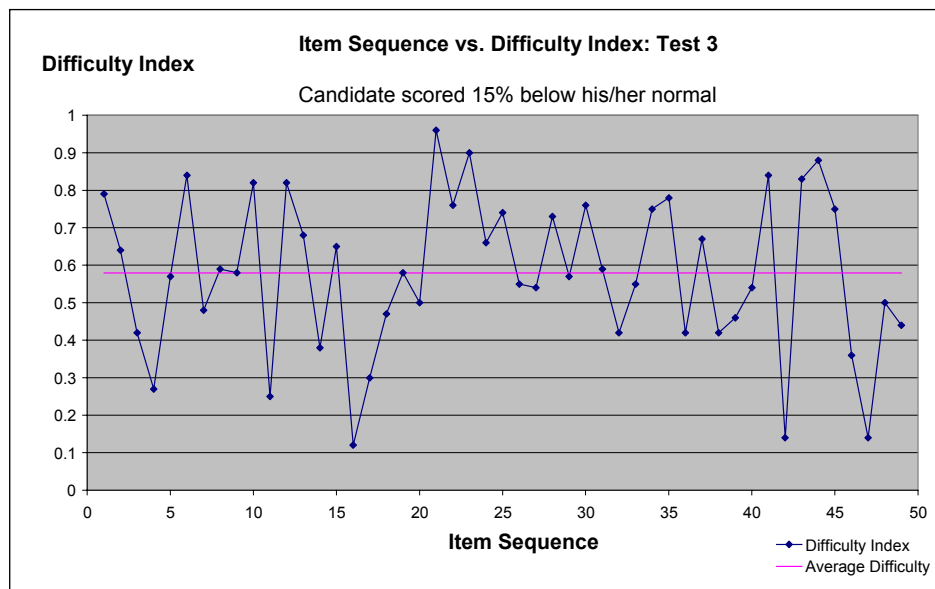


Figure 3. Example from Test 3

Figure 4 shows a candidate from the forth test who scored 15 percent below his/her normal. Of the first sixteen items, three were easy, four difficult, and nine of average difficulty. Again, this shows that the initial third of the test was of average difficulty, which is ideal for the middle portion of the test, but not for the initial portion of the test. Even closer inspection shows only six items above the average line and ten items on the difficult side below the average line. The initial items presented to this candidate clearly tended to be the more difficult items. This indicates to us that this student too could have been a victim of unfair assessment by having been randomly presented with difficult items early in this test.

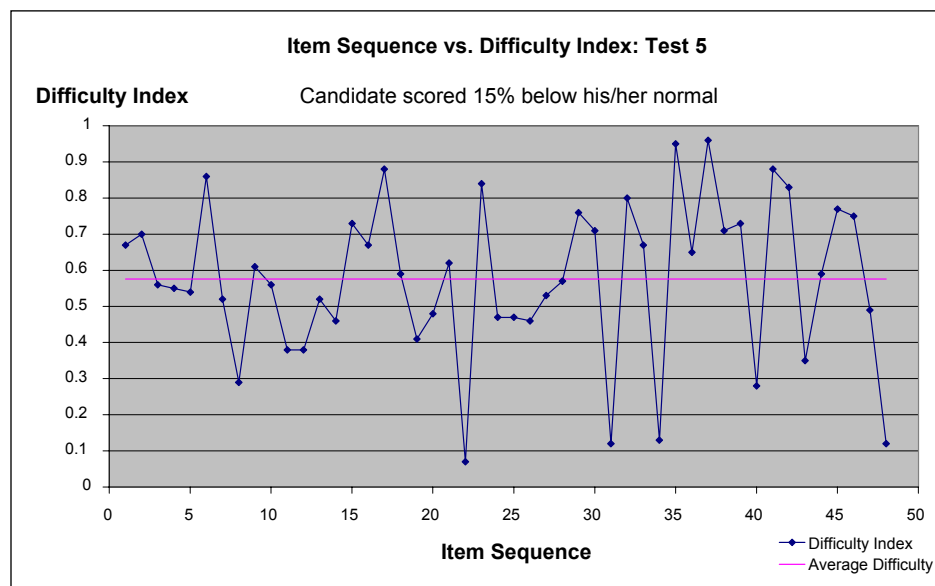


Figure 4. Example from Test 4

## Limitations

The sample is far too small for such a study, and therefore one can only conclude that an experimental study must be specifically designed, in which all the candidates in the experimental group are presented several difficult items consecutively, and all candidates in the control group are presented with several easy items consecutively, with the results of the two groups then compared.

Formal statistical analysis was conducted on the cleaned data, more in an effort to cover all the bases and not because any significant correlations were expected to be found for this ideographic investigation. As was expected, no statistically significant differences were found. However, the infrequency of the affected test-takers in the sample does cause the statistical software to generate a warning that “chi-square may not be a valid test.” Table 1 below illustrates the problem with analysing the data. For each of the tests, fewer than ten percent of the sample scores vary by at least 15 percent from the test-takers’ normal performance, calculated as explained earlier in this paper. In fact, the percentage drops below five for the last two tests. This is mainly due to the decision to use existing computer-based test data; however, a properly designed experimental research project would ensure the study becomes nomothetic, and hence statistical analysis becomes useful for the investigation of potential correlations.

*Table 1. Scores varying by at least 15% from normal*

	<b>Test 1</b>	<b>Test 2</b>	<b>Test 3</b>	<b>Test 5</b>
<b>Scoring lower</b>	2	5	3	2
<b>Scoring higher</b>	7	2	0	0
<b>Total</b>	9	7	3	2
<b>% of Sample</b>	<b>8.7%</b>	<b>6.7%</b>	<b>2.9%</b>	<b>1.9%</b>

The above limitations are explained further in terms of construct validity in Bugbee Jr. (1996), with regard to the equivalence of tests across modes. An error of measurement (error variance) is contrasted with systematic variance with respect to the mode of administration. The important distinction is in noting that systemic variance affects all test-takers equally, whereas “error variance or error of measurement is variation of errors due to chance” (Bugbee Jr., 1996, Specific Research Studies section, para. 12). This study is typical of error variance, and error variance in this context shows that either the variance is not due to the mode of administration, or that presenting difficult items early will likely cause as much increased anxiety in paper-based tests too. We are not contradicting this, but are concerned that the effect is more pronounced in computer-based test assessments due in part to the lesser convenience in this mode for navigating through the test items. One way around this dilemma is to design a test across modes that ensures that all candidates are presented with the difficult items early such that the assessment gets progressively easier. The research design ensures to some extent a systemic variance in that all test-takers may be affected equally. Care must be taken to account for those candidates with good test-taking skills and low levels of test anxiety who will probably attempt the items in a progressively more difficult sequence, influencing the reliability of the study.

## Conclusions and recommendations

Despite the limitations of this study, the findings are noteworthy because a potentially unfair testing practice has been identified and verbalised. Also, a potential gap in the literature can now be filled as researchers investigate the test-mode effect that, in this study, is referred to as the Item Randomisation Effect.

Unfairness in assessment is not an acceptable practice for test developers, users, or takers. The randomly affected test-taker is the one who suffers any consequences of this practice, yet it is within the powers of the test developers to ensure this won’t happen by programming algorithms in computer-based testing software. In addition, it is the responsibility of test users to ensure developers are made aware of this test mode effect. One test-taker affected by this unintended Russian roulette in assessment is one too many, as it would violate the APA guidelines on fair testing (APA, 2004).

Assuming this mode effect is found to cause some students to be disadvantaged in computer-based test assessments, we recommend that software vendors add a few lines of code to the randomising algorithm, such that this test mode effect is automatically prevented from occurring in future computer-based test assessments. This is easily achieved

once items have been used in a test, as each item can now be ranked according to its difficulty index, labeled, and stored in the item database. This could be useful for ensuring that randomising algorithms present items to students randomly while progressively increasing the difficulty of the items. Randomisation to ensure test security, yet progressively allowing items to become more difficult as the test items are presented to each test-taker, will prevent occurrence of the item randomisation effect.

It is hoped that this paper will inspire researchers to create studies specifically suited to obtain correlations that will confirm or refute the item randomisation effect introduced in this paper.

## Acknowledgements

Thank you to the University of Pretoria, Computer Testing Section, for their assistance and for supplying the test data; the Faculty of Veterinary Sciences for permission to use their data; and the Department of Statistics, particularly Mrs. Jaqui Sommerville, for the formal data analysis.

## References

- American Psychological Association. (1988). Rights and responsibilities of test takers: Guidelines and expectations. Retrieved January 16, 2006, from [www.apa.org/science/ttrr.html](http://www.apa.org/science/ttrr.html).
- American Psychological Association. (2004). Code of fair testing practices in education. Retrieved January 16, 2006, from [www.apa.org/science/fairtestcode.html](http://www.apa.org/science/fairtestcode.html)
- Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning: Methods and development* (3rd ed.). Boston: Allyn & Bacon.
- Bemelmans, K. J., Wolters, G., Zwinderman, K., ten Berge, J. M. F., & Goekoop, J. G. (2002). Evidence for two processes underlying the serial position curve of single- and multi-trial free recall in a heterogeneous group of psychiatric patients: A confirmatory factor analytic study. *Memory*, 10(2), 151.
- Bierenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education*, 53(6), 749–768.
- Billings, K. (2004). Online assessment: Perspectives of developers. *Media & Methods*, 40(4), 26.
- Black, S. (2005). Test anxiety. *American School Board Journal*, 192(6), 42.
- Bugbee Jr., A. C. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education*, 28(3), 282.
- Carlson, J. F., & Smith Harvey, V. (2004). Using computer-related technology for assessment activities: Ethical and professional practice issues for school psychologists. *Computers in Human Behavior*, 20(5), 645.
- Dutke, S., & Stöber, J. (2001). Test anxiety, working memory, and cognitive performance: Supportive effects of sequential demands. *Cognition & Emotion*, 15(3), 381.
- Ferguson, K. J., Kreiter, C. D., Peterson, M. W., Rowat, J. A., & Elliott, S. T. (2002). Is that your final answer? Relationship of changed answers to overall performance on a computer-based medical school course examination. *Teaching & Learning in Medicine*, 14(1), 20–24.
- Glenn, R. E. (2004). Teach kids test-taking tactics. *Education Digest*, 70(2), 61–63.
- Hancock, D. R. (2001). Effects of test anxiety and evaluative threat on students' achievement and motivation. *Journal of Educational Research*, 94(5), 284.
- Harding, R. (2001). What have examinations got to do with computers in education? *Journal of Computer Assisted Learning*, 17(3), 322.
- Hoff, D. J. (1999). Testing. *Education Week*, 19(2), 6.

- Lufi, D., Okasha, S., & Cohen, A. (2004). Test anxiety and its effect on the personality of students with learning disabilities. *Learning Disability Quarterly*, 27(3), 176.
- Pain, D. E., & Le Heron, J. L. (2003). WebCT and online assessment: The best thing since SOAP? *International Forum on Educational Technology & Society*, 6(2), 62–71.
- Reduce test anxiety to improve student performance. (2005). *Teaching Professor*, 19(9), 5.
- Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of Personality Assessment*, 81(2), 93.
- Russell, M., Goldberg, A., & O'Connor, K. (2003). Computer-based testing and validity: A look back into the future. *Assessment in Education: Principles, Policy & Practice*, 10(3), 279.
- Staber, J., & Pekrun, R. (2004). Advances in Test Anxiety Research. *Anxiety, Stress & Coping*, p. 205.
- Sternberg, R. J. (1998). Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science*, 26(1–2), 127–140.
- Supon, V. (2004). Implementing Strategies to Assist Test-Anxious Students. *Journal of Instructional Psychology*, 31(4), 292–295.
- Talmi, D., & Goshen-Gottstein, Y. (2006). The long-term recency effect in recognition memory. *Memory*, 14(4), 424.
- Tobias, S. (1985). Test anxiety: Interference, defective skills, and cognitive capacity. *Educational Psychologist*, 20(3), 135–142.
- Tseng, H.-M., Tiplady, B., Macleod, H. A., & Wright, P. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computer and paper assessment of mood and performance. *British Journal of Psychology*, 89(4), 599.
- Turner, S. M., DeMers, S. T., Fox, H. R., & Reed, G. M. (2001). APA's guidelines for test user qualifications. *American Psychologist*, 56(12), 1099.
- Use parent nights to improve student test-taking skills. (2004). *Curriculum Review*, 43(5), 6.
- Varughese, J. A. (2005). Testing, testing. *University Business*, 8(4), 59.
- Zenisky, A. L., & Sireci, S. G. (2002). Technological innovations in large-scale assessment. *Applied Measurement in Education*, 15(4), 326–337.

# Investigation of Interaction, Online Support, Course Structure and Flexibility as the Contributing Factors to Students' Satisfaction in an Online Certificate Program

**Erman Yukselturk and Zahide Yildirim**

Department of Computer Education and Instructional Technology, Middle East Technical University, Turkey // eyukselturk@gmail.com, ermany@metu.edu.tr // zahidey@metu.edu.tr

## ABSTRACT

The purpose of the study is to investigate some of the factors that contribute to satisfaction of participants in an online Information Technologies Certificate Program (ITCP). The program includes eight courses and lasts four semesters. The sample consisted of 30 participants who enrolled to the program and 8 instructors who gave the courses in this program in 2004-2005. The data was collected through both qualitative and quantitative methods. An online questionnaire was used to gather data on participants' satisfaction about the program, semi-structured interviews were conducted with both the participants and the instructors to analyze the factors that contribute to satisfaction in the program, and asynchronous and synchronous communication transcripts were examined to support the findings. Descriptive statistics were reported upon the participants' satisfaction in regard to learner-learner interaction, learner-instructor interaction, course structure, institutional support, and flexibility. A repeated measure analysis test was utilized to see the changes in the participants' satisfaction throughout this online program based on semester one, two, three, and four. The results of the study showed that even though the participants' overall satisfaction was generally positive, it decreased significantly toward the final semesters of the program. The findings highlighted some of the critical issues such as "learning community formation through interaction, well structured and guided project or problem-based group activities, considering instructional design and cognitive principles in the design of the program, and flexibility in time and selection of the courses" that should be taken into account in designing online programs.

## Keywords

Online learner satisfaction, Online interaction, Online support, Online course structure, Online flexibility

## Introduction

The recent emergence of information and communication technologies allows universities to offer distance education programs to meet the needs of nontraditional students. With the help of these technologies, the number of online courses and programs has increased drastically in the recent years. It was also expected that the number of online students will grow in the long-term (Allen & Seaman, 2004).

With the increasing number of online courses and programs, the number of institutions that provide them is also increasing. This causes competition among providers in the field of distance education (Tricker, Rangecroft, Long & Gilroy, 2001). There are factors that affect the teaching quality such as good teaching, clear goals, appropriate workload, appropriate assessment and emphasis on independence (Ramsden, 1991). Even though it is not the only factor that affects teaching quality, in this new competitive field, the focus has been shifted to learners' needs, expectations and satisfaction. Especially, learner satisfaction is more important than ever before in this new field (Roach & Lemasters, 2006). While promoting the quality of online programs in today's market, higher education institutions consider student satisfaction as one of the major principles (Moore, 2002; Moore & Kearsley, 2005). Feasely and Olgren (1998) mentioned that according to Kirkpatrick's (1998) four levels model of learning which consists of reaction, retention, application, and results, the learner's reaction to course material is categorized as the first level. In the part of reaction level, measuring learner satisfaction provides valuable information about the attentiveness of the student, the overall learning experience, and the effort exerted to learn. High level of learner satisfaction leads to lower attrition rates, an increase in learners' enrollments and motivation, and a more productive learning environment (Biner, Dean, & Mellinger, 1994; Schwitzer, Ancis, & Brown, 2001).

Student satisfaction is seen as one of the key variables in determining the success or failure of distance learners, courses, and programs in the literature. Therefore, there are many published studies on distance learner satisfaction. However, earlier studies tended to focus on learner satisfaction measured for once only at the beginning or at the end of the course with questionnaires, grades given on tests and other course assignments. Therefore, Sener and Humbert (2003) stated that more longitudinal studies are needed to investigate this issue. Phipps and Merisotis (1999) stated



that a major shortcoming of the distance learning research to date was the emphasis on student outcomes for individual courses rather than for whole academic programs. Roach and Lemasters (2006) mentioned that the researches in the literature were not specific to online professional programs.

To sum up, with the help of multiple sources that were used to collect data, such as questionnaires, student and instructor interviews, chat and discussion list transcripts, this study will help to answer questions related to offering entire online programs. Further, the results of the study help us avoid the pitfalls of a one-shot measure of student satisfaction, and provide a better understanding of changes in student satisfaction in a long period of time and in entire programs. At this stage, the theoretical framework about student satisfaction is provided to form a base for this study.

## **Learner Satisfaction**

Learner satisfaction in distance education depends on a number of factors (Sener & Humbert, 2003). In this study, satisfaction is defined by learner-reported feelings about interaction with instructors and peers, course structure, institutional support, and flexibility.

Interaction is an important part of learner satisfaction. Research suggests that both quality and quantity of interaction with the instructor and peers are much more crucial to the success of online courses and student satisfaction than that are in traditional courses (Woods, 2002). In a study by Fulford and Zhang (1993), the students' perception of interaction was the critical predictor of satisfaction in a distance-learning course. In another study conducted to determine predictors of student satisfaction with a fully interactive, multi-point real-time video teleconferencing and web-based course, Debourgh (1999) found that the factors that were related to interaction were critically important. These factors are related to promptness of answers to student questions, instructors' encouragement of participation, accessibility of instructors, and promptness of instructors' feedback on students' work. In addition, Swam (2001) analyzed satisfaction of university level online students (1406 students from 73 online courses) and stated that three factors contributed significantly to satisfaction level in online courses. These factors are contact with and feedback from the instructors, active discussion among students, and clarity in course design. This result is consistent with Moore's (1989) promotion of three types of interaction in distance education: learner-instructor, learner-learner, learner-content.

Course structure is another important factor that affects the learner satisfaction. A well-designed online course is beyond publishing text. In other words, designing an online course is not simply converting traditional course material to an online format. Course structure helps distance learners plan, organize and manage their learning activities. In a study by Hara and Kling (1999), dissatisfaction with the distance course is related to the lack of prompt feedback, technical difficulties, and ambiguous course instructions.

High quality of instructional and institutional support services also result in higher student satisfaction with the educational environment. Schwitzer, Ancis and Brown (2001) stated that interactive and engaging student services were critical factors for student satisfaction. Moore and Kearsley (2005) suggested that students might interact with specialists in various forms of student support in addition to interacting with instructors who help students learn the course contents. Student's satisfaction might also increase when on-site staffs are attentive to their progress and problems (Biner, Dean, & Mellinger, 1994).

In distance education programs, the learners are mainly adult with social, occupational and family obligations (Moore & Kearsley, 2005). Therefore, distance learners want to participate in the courses or programs whenever and wherever they need or want. In other words, the convenience and flexibility of distance program is another contributor to student satisfaction (Arbaugh, 2000). Maki, Maki, Patterson and Whittaker (2000) found that students enjoyed the flexibility of online learning environment and they perceived the convenience of the online course as a major benefit. Increasing learner control by providing flexibility and different choices within the instructional programs help strengthen learner satisfaction.

To summarize, student satisfaction has been given high attention in distance learning environment like in traditional learning environment in recent years. There are number of reasons for the attention given to learners' satisfaction in online courses. First, understanding the factors that affect the student satisfaction such as interaction with instructors and peers, course structure, institutional support, and flexibility will help course designers and teachers take the

necessary measures to increase satisfaction. Second, it is seen as an important measure of program outcomes and program quality. Third, satisfaction is influenced a lot by contextual factors, making it necessary to analyze online learning within its own context. The purpose of this study is to investigate some of the contributing factors such as interactions, course structure, institutional support and flexibility to satisfaction of the participants enrolled in the online Information Technologies Certificate Program. The following research questions guided this study:

- What is the participants' satisfaction in regard to learner-learner interaction, learner-instructor interaction, course structure, institutional support, and flexibility, based on semester one, two, three and four in online Information Technologies Certificate Program (ITCP)?
- Does the participants' satisfaction change throughout online Information Technologies Certificate Program (ITCP) based on semester one, two, three, and four?
- What are the factors that contribute to the participants' satisfaction in online Information Technologies Certificate Program (ITCP)?

## **Online Information Technologies Certificate Program (ITCP)**

Many higher education institutions have made the decision to offer online courses, certificate programs and full degree programs in recent years. One of these programs is online Information Technologies Certificate Program (ITCP). This program was the first that became valid in Turkey, it was developed at Middle East Technical University (METU) with the enterprise of METU Computer Engineering Department, technical support of METU Computer Center and collaboration of METU Continuing Education Center in 1998, and it is still active. It includes eight fundamental courses of Computer Engineering Department, and comprises of four semesters lasting nine months totally. The courses in the program are given by the instructors from Computer Engineering Department. The main aim of the online ITCP is to train the participants in the IT field to meet the demands in the field of computer technologies in Turkey. Furthermore, the online ITCP provides opportunities for the people who could not get education in information technologies or computer engineering, but interested and willing to improve themselves in this area and who are enthusiastic about making progress in their existing career. University students and people who graduated from 2 or 4 year university programs have been accepted to the programs. In addition, the participants are expected to be computer literate and to have competency in English at intermediate level (Isler, 1998a; 1998b).

The program provides online lecture notes, learning activities and visual aids, and each course has a textbook to follow. An instructor and an assistant are assigned for each course. In order to provide interaction between instructors and participants, and among participants, each course has an e-mail address, discussion list and chat sessions. At the end of each term, there are face-to-face sessions for each course within the campus of METU. For each course, at least three or four assignments are given to the participants during the semesters. At the end of each semester, there are traditional final examinations within the campus of the University. The participants' final grades are based on the final examinations, assignments, attendance to chat sessions and discussion lists. At the end of the program, graduates receive official certificate approved by the president of METU, the chair of the Computer Engineering Department and the president of the Continuing Education Center (Isler, 1998a; 1998b). The courses given in this program are as follows:

### **First Semester (lasting two months)**

- Computer Systems and Structures
- Introduction to Computer Programming with C

### **Second Semester (lasting two months)**

- Data Structure and Algorithms with C Operating Systems with Unix

### **Third Semester (lasting two months)**

- Software Engineering
- Database Management Systems

### **Fourth Semester (lasting two months)**

- Web Programming
- Software Development Project

## **Method**

### **Design of the Study**

This study is a mixed methods case study which is an in-depth study of a chosen event, activity, process of group using extensive data collection (Merriam, 1998). A case study approach is advantageous when “why” and “how” questions are being asked, and it is recommended when the researcher believes that the contextual conditions are highly relevant to the phenomenon under study (Yin, 1994). It means that case study method is useful to understand

particular situation, course, and program in depth, such as online certificate program. A program-level case study was used to determine how satisfied the online certificate program's participants were in general, and what factors contributed to their satisfaction in this study. A combination of quantitative and qualitative methods was used to collect relevant data in this case study. The satisfaction of participants about this program was examined in depth through interviews, questionnaire and online communication transcripts. The careful and purposeful combinations of different methods in social and behavioral research strengthen and deepen the analysis, and decrease the weaknesses of the study (Johnson and Turner, 2003).

## **Participants**

The study included 30 ITCP participants and eighth instructors at Middle East Technical University in Ankara, Turkey (October 2004 - June 2005). Originally, sixty two students were registered to the program, however this study included the students the ones who completed the program, got IT certificate, and were volunteer to participate in the study. All participants who registered to the program were computer literate and had an intermediate level of English. The number of male participants (N= 22) was greater than the number of female participants (N=8), and the participants' age ranged from 20 to 40 and above. The majority of the participants' ages were between 20 and 29 (N=21). The majority of the participants were university graduates and undergraduate students.

The eight instructors included in this study were the ones who gave the eight online courses in the program. The instructors were interviewed individually about the factors that affect or contribute to participants' satisfaction. The instructors were faculty members at the Department of Computer Engineering of the university. All of them have given online courses in this program for over six years, and they developed their own course materials.

## **Instrumentation**

An online questionnaire was used to collect data on participants' satisfaction about the program. The questionnaire was developed based on four questionnaires in the literature (Instructor and Course Evaluation System, ICES, (University of New Mexico, 2001), Distance and Open Learning Environment Scale, DOLES, (Jegede, Fraser and Curtin, 1995), Class Interaction, Structure and Support, CISS, (Johnson et al., 1999), and Web-Based Learning Environment Inventory, WEBLEI, (Chang and Fisher, 1999). It consists of five main subscales; (1) Learner-learner interaction (with 3 questions, the Cronbach Alpha value is 0.80); (2) Learner-instructor interaction (with 12 questions, the Cronbach Alpha value is 0.93); (3) Course structure (with 12 questions, the Cronbach Alpha value is 0.92); (4) Institutional support (with 12 questions, the Cronbach Alpha value is 0.88); and (5) Flexibility (with 3 questions, the Cronbach Alpha value is 0.59). It consisted of 38 five-point likert type items, and the overall Cronbach Alpha value of the questionnaire is 0.95 (Parlak, 2004).

In addition to the questionnaire, semi-structured interviews were conducted with the participants and the instructors to elicit additional information regarding participants' satisfaction. The interview schedules were developed around the central themes related to the components of the satisfaction (interaction with instructors and peers, course structure, institutional support, and flexibility), and consisted of 16 questions for the participants and 12 questions for the instructors. The interview schedules were examined for clarity by three experts in the field of instructional technology at the university. Additionally, pilot interviews were conducted with 5 participants and 2 instructors in the previous program, and based on the feedback gathered from the pilot interviews, the interview schedules were improved.

## **Data Collection and Analysis**

The Online Certificate Program comprises of four semesters lasting nine months totally. After completing each semester lasting two months, the participants were requested to fill out the online satisfaction questionnaire. The data gathered through the questionnaire were analyzed by descriptive and inferential statistics such as means and standard deviations, and repeated measures analysis and *t* test to examine the changes in the participants' satisfaction throughout the first, second, third and fourth semesters.

Semi-structured interviews were conducted with three participants and two instructors at the end of each consecutive semester. Until the end of the program, interviews were conducted with totally 12 participants and 8 instructors. Bogdan and Biklen (1998) stated that researchers are confident of receiving comparable data with semi-structured interviews. In order to represent variety among participants, the interviewers were selected from different participation levels (high, average and low) in online discussions, and from different field of experience. Before each interview took place, the participants and instructors were informed of the purpose of the interview. Each interview took about 20-30 minutes, and was tape-recorded with the permission of the participants. The data analysis was continuous and iterative throughout the data collection and the report writing. This analysis process went through in iterative cycles of examining the data, exploring similarities and differences among the participants, and searching for confirming and disconfirming evidence that would be incorporated into the conclusions (Merriam, 1998). Additionally, participants' online discussion logs were examined to find out the level of participation to the online communication in the program.

## Results

### The Participants' Satisfaction Level in the Program

Table 1 presents the participants' satisfaction in regard to learner-learner interaction, learner-instructor interaction, course structure, institutional support, and flexibility. The participants' overall satisfaction scores were  $M=3.5$ ;  $M=3.5$ ;  $M=3.4$ ; and  $M=3.2$  at consecutive semesters one, two, three and four indicating that majority of the participants agreed with the statements for the first, second and third semesters but they were neutral for the fourth semester. When overall satisfactions of the participants in regard to satisfaction themes were examined, learner-learner interaction had the lowest mean (2.9), and course structure had the highest mean (3.9). The results also indicated that the participants did not agree that the fourth semester's courses were flexible enough ( $M=2.3$ ).

Table 1. The Participants' Satisfaction in the Semesters of the Program

	Semester-1		Semester-2		Semester-3		Semester-4		Overall	
	M	SD	M	SD	M	SD	M	SD	M	SD
Learner-learner interaction	2.7	1.0	2.8	1.1	2.9	1.0	3.0	1.1	<b>2.9</b>	<b>1.1</b>
Learner-instructor interaction	3.6	0.7	3.7	0.9	3.6	1.0	3.5	1.1	<b>3.6</b>	<b>0.9</b>
Course structure	4.1	0.8	4.1	0.8	3.8	0.9	3.6	0.9	<b>3.9</b>	<b>0.9</b>
Institutional support	3.4	0.8	3.4	1.1	3.5	0.9	3.4	1.0	<b>3.4</b>	<b>1.0</b>
Flexibility	3.9	0.8	3.7	1.1	3.1	0.9	2.3	0.9	<b>3.3</b>	<b>0.9</b>
<b>Overall Satisfaction in each semester</b>	<b>3.5</b>	<b>0.8</b>	<b>3.5</b>	<b>1.0</b>	<b>3.4</b>	<b>0.9</b>	<b>3.2</b>	<b>1.0</b>	<b>3.4</b>	<b>0.9</b>

M: Mean, SD: Standard Deviation

### The Changes (Variations) in the Participants' Satisfaction throughout the Program

In order to answer whether participants' satisfaction changes throughout the semesters of the program, a repeated measure analysis and *t* test was employed at a significance level of .05.

The repeated measure analysis indicated that there was no significant difference among the means of four semesters in regard to *learner-learner interaction*,  $F(3, 87) = 0.82$   $p=0.48$ ; *learner-instructor interaction*,  $F(3, 87) = 0.65$   $p=0.58$  and the *institutional support*,  $F(3, 87) = 0.28$   $p=0.83$ .

In regard to satisfaction on *course structure*, the repeated measures analysis showed that there was a significant difference among the means of four semesters,  $F(3, 87) = 7.2$   $p=0.00$ . Course structure measures decreased especially during the last two semesters. In order to find out the means differ from each other at a significant level, a paired-samples *t* test was used. The *t* test results indicated that there was a significant mean difference between the first and the fourth semesters ( $p=0.03$ ), the second and the third semesters ( $p=0.013$ ), the second and the fourth semesters ( $p=0.01$ ), and the third and fourth semesters ( $p=0.023$ ) on course structure satisfaction.

The repeated measures analysis also indicated that there was a significant difference among the means of four semesters for the *flexibility* measures,  $F(3, 87) = 26.86$   $p=0.000$ . Flexibility measures decreased gradually

throughout the semesters. The paired-samples  $t$  test was used to find out the means differ from each other. The results showed that there was a significant mean difference between the first and the third semesters ( $p=0.001$ ), the first and the fourth semesters ( $p=0.001$ ), the second and the third semesters ( $p=0.001$ ), the second and the fourth semesters ( $p=0.000$ ), and the third and the fourth semesters ( $p=0.000$ ) in regard to satisfaction in flexibility.

The results of the study showed that there was a statistically significant decrease,  $F(3, 87) = 5.35$ ,  $p = 0.002$  in *overall satisfaction* of the participants through out the program. As it shown in Figure 1, overall satisfaction decreased gradually from the first to the fourth semesters. The paired-samples  $t$  test results showed that there was a significant mean difference between the first and the fourth semesters ( $p=0.01$ ), the second and the fourth semesters ( $p=0.008$ ), and the third and the fourth semesters ( $p=0.005$ ) in regard to overall satisfaction of the participants.

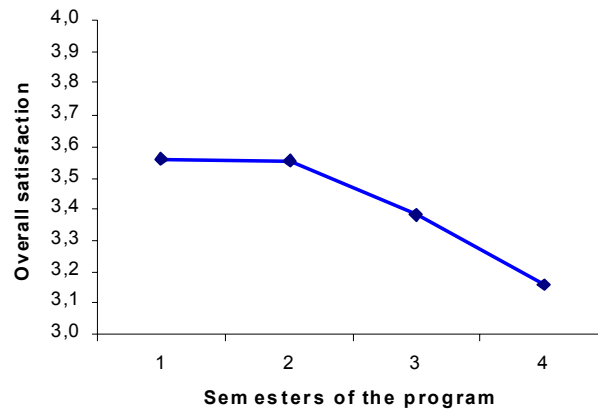


Figure 1. Participants' Overall Satisfaction throughout the Semesters

## Interaction

### *Usage of CMC Tools*

In this online certificate program, mainly asynchronous (e.g. discussion list, e-mail) and synchronous (e.g. chat sessions) communication tools were used to facilitate interaction among the participants, and between the participants and instructors.

The participants and instructors thought that learning with interaction through CMC tools was one of the major benefits of online learning when it is compared to learning from texts by themselves. While the participants used e-mail to ask about specific issues that were primarily related to their personal problems, the instructors used e-mail when the content of the e-mail was not related to all participants. Furthermore, almost all instructors and participants expressed that the course discussion lists were the most useful and preferable tool in this online program, and they used them mostly for interaction and communication in the program. They wrote messages and received answers in a short time in the discussion lists. The participants mentioned that they could write longer messages in the discussion list compared to the chat sessions. Therefore, they could discuss more topics in a detailed way. According to the instructors, another major advantage of discussion lists was that participants could see others' problems and solutions which they might also have faced. The instructors generally posted messages in the discussion list to announce course related issues such as, assignments, exams. Moreover, with the chat sessions, instructors and participants could ask questions and get immediate answers in real time. Original and natural discussion environment similar to the face-to-face environment could be created in this way. In brief, asynchronous and synchronous communications were used complementarily to prove more beneficial in the online program since instructors thought that both of them have different weaknesses and strengths.

### *Interaction among the Participants*

The findings of interviews with instructors and participants, and the findings of chat session and discussion list transcripts supported the results of the questionnaire on interaction in the program. Almost all participants and instructors expressed that the interaction in the program, especially among the participants, was not enough in all semesters. Table 2 shows that the participants' participation to the online communication in the program was insufficient. Even though the number of messages sent by participants, some of which was related to meeting and introducing themselves, was high in the first semester, it was very low at the chat sessions and discussion lists in the last three semesters. Although the number of messages was higher in chat session than that was in discussion list, the number of participants in the discussion list was greater than that was in the chat sessions.

*Table 2. The Participants' and the Instructors' Participation to Online Communication*

	Semester-1		Semester-2		Semester-3		Semester-4		Overall	
	M	SD	M	SD	M	SD	M	SD	M	SD
# of attended participants in CS	17.5	6.0	7.1	3.6	5.6	0.7	5.9	2.1	9.0	3.1
Total # of messages sent by participants in CS	525	231.1	74.4	34	104	58.2	111	66.8	203.6	97.6
Total # of messages sent by instructors in CS	276	56.8	112	38.6	144	45	140	45	168	46.4
<b>Total messages in CS</b>	<b>804.8</b>	<b>247.9</b>	<b>186</b>	<b>38.2</b>	<b>248</b>	<b>97</b>	<b>251</b>	<b>89</b>	<b>372.5</b>	<b>118</b>
# of attended participants in DL	29	1.4	12	1.4	12	0.0	13.5	4.9	16.6	1.9
Total # of messages sent by participants in DL	174.5	0.7	37	1.4	38	9.8	76	35.3	81.36	11.8
# of messages sent by instructors in DL	112	7.07	47.5	12.0	43	28.2	65	7.1	66.86	13.6
<b>Total messages in DL</b>	<b>286.5</b>	<b>7.7</b>	<b>84.5</b>	<b>13.4</b>	<b>81</b>	<b>38.1</b>	<b>141</b>	<b>42.4</b>	<b>148.3</b>	<b>25.4</b>

CS: Chat Sessions, DL: Discussion Lists, M: Mean, SD: Standard Deviation

Even though the interaction among the participants was weak in the program, interview results showed that participants wanted to interact among themselves during the program. For instance, five participants thought that the participants in the program could help each other by using communication tools. Although the participants were eager, the majority of them did not get help from and interact with others often during the program. Moreover, participants mentioned that they sometimes attempted to interact with others in the program. For example, they tried to interact and meet with others during the face-to-face courses at the end of the semesters. Also, some participants submitted their ICQ numbers to ICQ club web sites in the program, and two participants attended program chat rooms two or three times randomly, but they did not encounter with the other participants. It can be concluded that although some participants tried to interact with others, they were not much successful to increase interaction with other participants.

### *Interaction with the Instructor*

In contrast to the level of interaction among participants, most participants agreed that they were satisfied with the interaction with the instructors in the program. For example, six participants mentioned that the instructors replied to their messages quickly, and they were quite satisfied with timely feedback. The findings of communication tool transcripts showed in Table 2 supported the participants' thought. For instance, overall number of messages sent by the instructors in both chat session and discussion lists was more than that of participants when the number of students was considered. According to the instructors, participants generally interacted with them about the requirements of the courses with which they had problems. They needed help from the instructors when they did not understand the course notes, exercises, and assignments.

### *The Reasons for Low Level of Interaction*

The participants and the instructors mentioned that there were several reasons for the low level interaction, especially among the participants. The first reason was due to the nature of the program that it was an online program.

Although there were face-to-face and chat sessions for each course in the program, they did not meet face-to-face continuously. They generally attended course web sites asynchronously; therefore, they generally were alone during their learning process. Moreover, they stated that the progress of interaction among the participants over the Internet took more time compared to face-to-face interaction.

The second reason was as stated by the participants and instructors that almost all participants had different responsibilities, and had various occupations in their life. They were busy, their life was dense, and their mind was full. Further, they stated that their background, previous knowledge levels, ages and occupation were different. The participants were over 18 and mostly set a different direction in their lives. This might influence the low level of interaction with others.

As the third reason, instructors mentioned that not studying regularly and difficulty in learning the subject hindered participants' interactions. Similarly, five participants mentioned that if they did not study regularly, they could not interact successfully in the courses. Two participants indicated that some courses were too difficult for them and they could not understand the topics. Therefore, they could not know what to ask or how to ask questions.

As the fourth reason, participants who had more knowledge wanted to ask more and complex questions, and they wanted to be dominant and more active; therefore, some participants did not attend the discussions. Similarly, the instructors stated that some participants might sometimes hesitate to attend discussions in order not to ask unrelated or silly questions.

Fifth, three participants stated that there were not enough activities to have interaction among the participants in the courses, so they could not find common topics to talk to each other and they might not improve their interaction.

Sixth, three instructors and two participants mentioned that some participants preferred to study alone. They did not need to interact with others much. Therefore, they only read communication log files. Also, some participants generally preferred to communicate with others who had common backgrounds or who knew each other before the program.

## **Course Structure**

Instructors stated that they tried to give eight basic computer engineering courses to the participants in this online program. All of them expressed that designing and giving these online courses properly was not easy job for them when compared to traditional one. They agreed that it was troublesome process for them to design online course materials especially at the first time. They knew that online education placed all major responsibilities on the students; therefore, they tried to create a more effective learning environment.

## *Assignments*

The results showed that most participants were satisfied with the contents of the courses that included various materials (e.g. exercises, assignments and reference books). Eight participants highlighted the importance of exercises and assignments by stating that exercises and examples in the courses helped them practice and learn the topics. With the help of preparing weekly assignments, they studied the course notes regularly, so that they could already be prepared for the exams.

## *Content*

Even though there were mostly positive thoughts about the course structure, some participants found deficiencies especially in the last two semesters' courses. Four participants mentioned that these courses consisted of too much general information, and examples and exercises were not sufficient to understand the topics. Five participants indicated that courses, especially during the third semester, contained a large amount of conceptual information. They could not understand them easily. Participants thought that even though they studied the course notes, they had to improve themselves more to understand the content. They had to read the textbooks and investigate other web

sites in the Internet to understand these concepts. These brought extra load and difficulties for the participants while they were studying.

Parallel to the participants, the instructors highlighted the major difficulty student faced while they were studying the theoretical and abstract course materials. They indicated that these materials were not easily understandable by the students just by reading especially when they face at the first time. At the same time, the instructors thought that online learners should be self-regulated. They mentioned that they tried to use self-study methods in their courses. They expected that the participants should learn the course topics not only through the course web pages but also through asking questions, searching, and investigating the provided resources.

### *Duration of the Program*

Another problem stated by the participants was about the duration of the program. They indicated that the program had a heavy and intensive curriculum since eight different courses were given in nine months. They could not grasp all the course topics in this limited amount of time. Also, a problem stated was the increased difficulty level of the courses from one semester to the next. Participants were expected to learn more than they could achieve in one semester. Likewise, the instructors agreed that this program had a heavy curriculum. Eight courses were given in the limited time and course contents were not easy to learn immediately. Therefore, they thought that the participants should study more than they do in traditional education.

### *Exercises and Examples*

Another concern of the participants was about the examples given in the courses. They thought that some examples were not applicable and practicable, and they could not apply what they learned. Participants suggested that course notes including examples, exercises and conceptual information should support online learning. These notes should be visual, clear and understandable to all.

The instructors also agreed that there were some deficiencies in the course structures, and they could be improved. Some course topics, especially the topics that were not easily comprehensible could be designed with the help of rich materials (i.e. interactive examples, multimedia applications). Further, they thought that course notes were updated regularly, some of them were changed and more current topics and contents were added. Additionally, the instructors indicated that the participants generally were satisfied with the courses when they implemented what they learnt to the real life immediately. Therefore, they tried to provide exercises to the participants in the course notes. However, these might not be sufficient for some students in certain courses.

## **Support**

The participants mentioned that this program provided many learner support activities to the participants during their education. They pointed out that while the registration period, they were informed about all the necessary issues and procedures. Throughout the orientation program, the course instructors met with them, and provided information about the program's properties. The participants also indicated that if they could not attend to face-to-face exams at the end of the semester, two more exams for each course were given at the end of the subsequent semesters. Additionally, they were pleased with the surveys given to them at the end of each semester for program and course evaluation.

Almost all instructors thought that the support given to the participants with one instructor and one or two assistants for each course was enough. They stated that they supported the participants with instructional activities during their education in the program, such as, face to face meetings at the beginning and at the end of the semesters, interaction through computer mediated communication tools (i.e. chat sessions, discussion lists), exercises, assignments and exams. Similarly, most participants agreed that they generally got response to their questions during the program. All kinds of problems, i.e. administrative, educational, communicational were tried to be solved by the program coordinators and instructors.



Even though a general consensus about the support system in the program was enough, the participants and instructors thought that it could be improved. They indicated that it might be helpful when participants were guided by the instructors individually during each course. Additionally, three participants mentioned that they sometimes needed immediate feedback while studying the course materials. If their questions were not solved at that time, they sometimes could not continue to study their courses. In addition, four participants mentioned that they needed more synchronous activities in the program. For example, number of chat sessions in each course might be increased; therefore, participants could attend suitable sessions.

## **Flexibility**

All instructors and participants mentioned that online education brought many advantages and disadvantages to both students and instructors when compared to traditional one. The major advantages of online education were providing much flexibility while taking or giving courses. For instance, the major flexibility of the program for almost all participants was that they were not obligated to be at a place or a school psychically to take the courses. They could study whenever and wherever they desired in this Internet-based program. Additionally, this program helped them learn eight basic computer engineering courses systematically.

## *Responsibilities for the Students*

This program brought some difficulties to the participants as well. For instance, most instructors and participants mentioned that this education placed all the responsibilities on the participants. They learned the content through the course materials and tried to apply them to new situations. Some instructors stated that several participants were not used to learning with this method, and it might not be proper for all participants.

The findings indicated that participants exerted too much effort to carry out their program related requirements. Almost all participants complained that they could not spend enough time to the courses due to their other responsibilities, i.e. job, family. Some participants stated that they sometimes sacrificed their special things to study for their courses (i.e. their hobbies).

Moreover, the instructors emphasized that participants had to study the course materials regularly and ask for help from peers or instructors when they needed. If they did not study the courses and perform their requirements regularly, they could not compensate them. They might stay behind all topics and might break off easily. After some time, the subjects to be studied were accumulated; therefore, they might not catch the others and loose their motivation to the courses. In such cases, participants might even leave the program.

## *Lack of Elective Courses*

The participants wanted more flexibility over the semesters in the program. For example, the aim of the program is to give eight basic computer engineering courses to the participants. However, some participants expressed that rather than taking all of the courses, they liked to take the courses they wanted. They also suggested that some elective courses should be added to the program.

## **Discussion**

Learner's satisfaction of online learning is regarded as positive in the literature (i.e. Allen & Seaman, 2004). A meta-analysis about studies of comparison between distance education and traditional methods showed that students find distance education as satisfactory as traditional classroom (see Allen, Bourhis, Burrell, and Mabry, 2002). In this study, the participants' satisfaction with online learning was generally positive. However, it decreased over the semesters of the program. The major factors considered in this study and affecting the satisfaction were interaction with instructors and peers, course structure, institutional support, and flexibility.

Interaction through communication tools seems to be one of the most influential features of online courses (Swan, 2001). The results of this study supported this idea. Asynchronous (e.g. discussion list) and synchronous (e.g. chat sessions) communication tools were used mainly for interaction among participants, and between participants and instructors in this online certificate program. According to the results, participants were generally satisfied with the interaction with their instructors in all semesters of the program. They were pleased for the effort of the instructors, and they could reach them easily. However, interaction among learners was not strong enough in this study. The lowest mean score factors affecting satisfaction was related to learner-learner interaction in the program. Similarly, the lack of interaction causes lack of attentiveness and lower satisfaction level with the learning experience in the literature (i.e. Bouhnik & Marcus, 2006; Daugherty & Funke, 1998; Swan, 2001).

The results demonstrated that there were number of reasons for low level interaction in the program. The reasons of low level interaction included having different responsibilities and various occupations, lack of time, interacting only with participants with common background or preferring to study alone, not enough possible interaction in Internet-based education environments, progress of interaction over the Internet was taking more time, not having enough interactive activities in the program, not studying course topics regularly, having little knowledge about the topics, hesitancy about writing messages, and participants' dominance in the discussions. Additionally, the participants could not create learning communities in the program. The lack of social interaction while learning in online environment resulted in the low levels of satisfaction. These issues related to the reasons of low level interaction were also stated by Dennen (2005), Hara and Kling (1999) and Northrup (2001) in their studies.

Participants' thoughts about the course structure were generally positive in the program. The basic computer engineering courses in the program were selected properly and their topics were arranged from the beginner to the advanced level. Participants were satisfied with the courses materials, e.g. examples, exercises, homeworks, reference books. Although participants' thoughts about the course structure were generally positive, their satisfaction decreased significantly over the semesters of the program, especially in the last two semesters. The interview results showed that some deficiencies such as lack of visual elements and richness in the materials in the courses affected the participants' satisfaction negatively. Additionally, loaded and intensive curriculum, and lack of elective courses in the program were the important points made by the participants.

Course structure and design appeared to be some of the key elements that affected student satisfaction along interaction. The results showed that even though some courses in the program were designed sufficiently, the structure of the whole program and design of the courses needed more attention. There was no one strategy for all courses and different strategies might be used for each course and semester (Northrup, 2001). As McPherson & Nunes (2004) indicated educators and designers must interact, and invest more time and effort on the analysis of learners, content, context, application, technologies, and curriculum to design for an effective online program (McPherson & Nunes, 2004).

Based on the research studies in distance education, Rumble (2000) summarized that learners without support delay program completion; contact between students and institution is beneficial; and advice given during enrollment affects later performance. In this program, the program coordinators and instructors tried to solve participants' problems when they arose. Participants were guided during the registrations and orientation program in this certificate program. Although the participants were satisfied with institutional support, the results showed that some participants desire more feedback while taking courses. They wanted to get individual and timely feedback about their progress. Thurmond et al. (2002) stated that instructors teaching online courses need to plan their schedules carefully for student evaluations and feedback activities since timely feedback from instructor contributed significantly to students' satisfaction. In line with this statement, the participants suggested more synchronous activities and more feedback.

The participants and instructors in this study agreed that this Internet-based certificate program provided some flexibility to them. For example, participants could study whenever and wherever they wanted. On the other hand, participants' satisfaction about flexibility of the program decreased significantly in the program. Over the semesters of the program, participants realized that this program brought some difficulties to them, such as responsibility of learning is on students, lack of time to spare for the program due to other responsibilities, lack of flexibility in course selection and the time of synchronous communication, and intense curriculum. These findings indicated that being ready and self-disciplined for such education are some of the characteristics that online students should possess.

Additionally, flexibility in the course selection, time of the online communication and extended amount of time for the program were the demands of the online learners.

These results emphasized the importance of responsibilities of online learners. In the literature, researchers stated that online learning placed more responsibilities on learners than traditional face-to-face learning did (Moore & Kearsley, 2005). It means that a different learning strategy, self-regulated learning, is necessary for online learning to be effective. Self-regulated learning requires changing roles of students from passive learners to active learners. However, the results of the study showed that participants might not be ready or adapted to this type of learning since, active learners or self-regulated learners select, organize, create advantageous learning environments for themselves and plan and control the form and amount of their own instruction for their academic achievement (Zimmerman, 2002).

As a summary, this study investigated some of the factors that affect satisfaction of participants enrolled in the online certificate program. The major factors that affected the participants' satisfaction and their changes over the semesters of the program and issues related to each factor were summarized in Table 3.

*Table 3. The Summary of Findings about Participants' Satisfaction in the Program*

Factors affecting satisfaction	Changes in the factors over the semesters	Issues related to each factor
Interaction with peers	Increase (not significantly)	<ul style="list-style-type: none"> <li>• participant's characteristics (i.e. lack of time, not studying course topics regularly)</li> <li>• internet's characteristics (i.e. progress of interaction over the Internet taking more time)</li> <li>• program's characteristics (i.e. not having enough interactive activities)</li> </ul>
Interaction with instructors	Almost the same	<ul style="list-style-type: none"> <li>• giving timely feedback</li> <li>• using both asynchronous and synchronous tools complementarily</li> </ul>
Course structure	Decrease (significantly)	<ul style="list-style-type: none"> <li>• designing course materials (i.e. examples, exercises, multimedia applications)</li> <li>• designing semesters and curriculum of program (i.e. numbers of courses, difficulties among semesters )</li> </ul>
Institutional support	Almost the same	<ul style="list-style-type: none"> <li>• designing institutional support activities (i.e. registrations and orientation program)</li> <li>• dealing with student problems regularly</li> <li>• individual and timely feedback</li> </ul>
Flexibility	Decrease (significantly)	<ul style="list-style-type: none"> <li>• having self-regulated learner skills</li> <li>• providing choices for the time of synchronous activities</li> <li>• providing enough time to fulfill program obligations</li> <li>• providing elective courses</li> </ul>

## Conclusions

Although it is not the only factor, satisfaction of the students is one of the important indicators of the success and quality of online programs. Therefore, this research analyzed a number of factors contributing to the learner's satisfaction for designing more effective online programs. In addition, as online education evolves, there is a number of rapidly growing factors that affect the learner satisfaction. Thus, it is essential to investigate what contributes to students' satisfaction in online education for a better design.

The findings of the study emphasized several critical issues for a better design in online programs. It can be stated that learning community formation and involvement plays a major role on student interaction. As Woods (2002) stated both quality and quantity of interaction with the instructor and peers are crucial for student satisfaction. In order to facilitate interaction among students, well structured instructional activities like projects or problems can be

designed and assigned to the groups, and guided by the instructors throughout the activities. The second finding was the necessity of giving importance to use of instructional design and cognitive principles in designing program courses and instructional materials. Arbaugh (2000) highlighted the importance of convenience and flexibility of distance program for student satisfaction. The findings of this study points out same direction. The third finding was the importance of having flexible online program and being self-regulated learner to benefit from such program. Together with this, it is also important to provide flexible time period (e.g. one course for each semester or time extension for each semester, choices for synchronous online sessions) for the participants since they have some other obligations to fulfill. Additionally, it would be convenient to provide elective courses to address different needs of the participants and make the program more flexible for the learners. Even though the findings of this study may help higher education institutions and online education designers consider some of the potential factors that affect students' satisfaction in entire online programs, one needs to be cautious in generalizing these findings to other online programs since the sample of this study was small. Therefore, further studies are needed to explore the factors that affect students' satisfaction in online learning environment in different subject areas, with larger student groups with different profiles. In addition, the relationship between the satisfaction and students' profiles such as online readiness, online technology competence, locus of control and self efficacy can be investigated.

## References

- Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *The American Journal of Distance Education*, 16 (2), 83-97.
- Allen, E., & Seaman, J. (2004). *Entering the mainstream: The quality and extent of online education in the United States, 2003 and 2004*, Needham, MA.: Sloan Center for Online Learning in Education (SCOLE).
- Arbaugh, J. B. (2000). Virtual classroom characteristics and student satisfaction with Internetbased MBA courses. *Journal of Management Education*, 24 (1), 32-54.
- Biner, P.M., Dean, R.S., & Mellinger, A.E. (1994). Factors underlying distance learner satisfaction with televised college-level courses. *The American Journal of Distance Education*, 8 (1), 60-71.
- Bogdan R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods*, Boston: Allyn and Bacon.
- Bouhnik, D., & Marcus, T. (2006). Interaction in distance-learning courses. *Journal of the American Society for Information Science and Technology*, 57 (3), 299-305
- Chang, V., & Fisher, D. (1999). *Students' perception of the efficacy of web-based learning environment: The emergence of a new learning instrument*, retrieved May 10, 2008 from <http://www.herdsa.org.au/wp-content/uploads/conference/1999/pdf/Chang.PDF>.
- Daugherty, M., & Funke, B.L. (1998). University faculty and student perceptions of web-based instruction. *Journal of Distance Education*, 13 (1), 21-39.
- DeBourgh, G. A. (1999). Technology is the tool, teaching is the task: Student satisfaction in distance learning. *Paper presented at the Society for Information Technology & Teacher Education International Conference*, February 28 – March 4, 1999, San Antonio, TX, USA.
- Dennen, V. P. (2005). From message posting to learning dialogues: Factors affecting learner participation in asynchronous discussion. *Distance Education*, 25 (1), 127- 148
- Feasely, C. & Olgren, C. (1998). *Evaluation for distance education*, Madison, WI: University of Wisconsin-Madison.
- Fulford, C. P., & Zhang, S. (1993). Perceptions of interaction: The critical predictor in distance education. *The American Journal of Distance Education*, 7 (3), 8-21.
- Hara, N. & Kling, R. (1999). Student frustrations with a web-based distance education course. *First Monday*, 4 (12), retrieved April 5, 2008 from [www.firstmonday.org/issues/issue4\\_12/hara/index.html](http://www.firstmonday.org/issues/issue4_12/hara/index.html).

- Isler, V. (1998a). Sanal Universite: Dusunceler ve Deneyimler. *Paper presented at Bilgi Teknolojileri Isiginda Egitim (BTIE)*, TBD and IEEE Turkiye Bolumu, May 22-24, 1998, Ankara, Turkey.
- Isler, V. (1998b) Internet Dayali Asenkron Egitim ile Bilgi Teknolojileri Sertifika. *Paper presented at YA/EM'98: Yoneylem Arastirma Konferansi*, June 25-26, 1998, Ankara, Turkey.
- Jegede, O.J., Fraser, B., & Curtin, D.F. (1995). The development and validation of a distance and open learning environment scale. *Educational Technology Research and Development*, 43 (1), 90-94.
- Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (1999). Comparative analysis of online vs. face-to-face instruction. In P. de Bra & J. Leggett (Eds.), *Proceedings of WebNet 99-World Conference on WWW and Internet*, Chesapeake, VA: AACE, 581-586.
- Johnson, R. B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori, and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*, Thousand Oaks, CA: Sage, 297-319.
- Kirkpatrick, D. (1998). *Evaluating training programs: The four levels*, San Francisco: Berrett-Koehler.
- Maki, R. H., Maki, W. S., Patterson, M., & Whittaker, P. D. (2000). Evaluation of a Web-based introductory psychology course: I. Learning and satisfaction in on-line versus lecture course. *Behavior Research Methods, Instruments, and Computers*, 32 (2), 230-239.
- McPherson, M. A., & Nunes, J. M. (2004) *Developing Innovation in Online Learning: An Action Research Framework*, London: RoutledgeFalmer.
- Merriam, S.B. (1998). *Qualitative Research and Case Study Applications in Education*, San Francisco: Jossey-Bass.
- Moore, M. (1989). Three types of interaction. *The American Journal of Distance Education*, 3 (2), 1-6.
- Moore, J. (2002). *Elements of Quality: The Sloan-C Framework*, Needham, MA: Sloan Center for Online Education.
- Moore, M. G., & Kearsley, G. (2005). *Distance education: A systems view (2<sup>nd</sup> Ed.)*, Belmont, CA: Wadsworth.
- Northrup, P. T. (2001). A framework for designing interactivity into web-based instruction. *Educational Technology*, 41 (2), 31-39.
- Phipps, R., & Merisotis, J. (1999). *What's the difference: A review of contemporary research on the effectiveness of distance learning in higher education*, Washington, DC: The Institute for Higher Education Policy.
- Parlak, O. (2004). *Student Satisfaction Scale on Internet Based Distance Education*, MS Thesis, Graduate School of Educational Science, Ankara University, Ankara, Turkey.
- Roach, V., & Lemasters, L. (2006), Satisfaction with online learning: A comparative descriptive study. *Journal of Interactive Online Learning*, 5 (3), 317-332.
- Ramsden, P. (1991) A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16 (2), 29-150.
- Rumble, G. (2000). Student support in distance education in the 21<sup>st</sup> century: Learning from service management. *Distance Education*, 21 (2), 216-235.
- Schwitzer, A., Ancis, J., & Brown, N. (2001). *Promoting student learning and student development at a distance: Student affairs principles and practices for televised instruction and other forms of distance learning (2<sup>nd</sup> Ed.)*, Lanham, MD: American College Personnel Association: University Press of America.
- Sener, J. & Humbert, J. (2003) Student satisfaction with online learning: an expanding universe. In J. Bourne & J. C. Moore (Eds.), *Elements of Quality Online Education: Practice and Direction*, Needham, MA: Sloan Center for Online Education, 245-260.
- Swan, K. (2001). Virtual interaction: Design factors affecting student satisfaction and perceived learning in asynchronous online courses. *Distance Education*, 22 (2), 306-331.
- Thurmond, V. A., Wambach, K., Connors, H. R., & Frey, B. B. (2002). Evaluation of student satisfaction: Determining the impact of a web-based environment by controlling for student characteristics. *The American journal of distance education*, 16 (3), 169-189.

Tricker, T., Rangecroft, M., Long, P., & Gilroy, P. (2001). Evaluating distance education courses: the student perception. *Assessment & Evaluation in Higher Education*, 26 (2), 165-177.

Woods, R.H. Jr. (2002). How much communication is enough in online courses? Exploring the relationship between frequency of instructor-initiated personal email and learners's perceptions of and participation in online learning. *International Journal of Instructional Media*, 29 (4), 377-394.

University of New Mexico (2001). *Complete ICES Forms*, retrieved May 28, 2008 from <http://its.unm.edu/ices/forms.html>.

Yin, R. (1994). *Case study research: Design and methods* (2<sup>nd</sup> Ed.), Beverly Hills, CA: Sage.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41 (2), 64-70.

## A Two-Dimension Process in Explaining Learners' Collaborative Behaviors in CSCL

**Jung-Lung Hsu, Huey-Wen Chou**

Department of Information Management, National Central University, Taiwan // srl666@gmail.com // hwchou@mgt.ncu.edu.tw

**Wu-Yuin Hwang**

Computer Science and Information Engineering, National Central University, Taiwan // wyhwang@cc.ncu.edu.tw

**Shyan-Bin Chou**

Graduate School of Design, National Taiwan Normal University, Taiwan // chou@ntnu.edu.tw

### ABSTRACT

Computer supported collaborative learning (CSCL) has captured many educators and researchers to contribute their efforts on this domain. This study proposed a two-dimension concept to explain learners' collaboration behaviors in a CSCL laboratory setting. A two-dimension process, namely perceptual dimensions and supportive dimensions, is useful to explain why learners' are willing to collaborate online. One hundred undergraduate students were divided into 20 work groups, each group consisting of 5 members. The result showed different group composition of extraverts will influence learners' perceptual dimension of group context variables, which in turn will be associated with supportive dimension of CSCL process. We conclude that educators and CSCL developers should focus not only on the technical factors of system development, but also on learners' psychological factors because of its impacts on the latter process-supportive dimension which especially stands for learners' collective behaviors of collaboration.

### Keywords

Knowledge sharing, Workload sharing, Collaborative process, Group norms, Task conflict

### Introduction

Computer supported collaborative learning (hereafter CSCL), as an emerging field in educational research, has captured many educators and researchers to contribute their efforts on this domain. Its anywhere - anytime characteristics and great potential to support interactive group learning make it one of the most promising next generation of educational tools. There is already an ample literature of research reporting its positive impact on learning (Altinay & Paraskevas, 2007; Ellis, 2001; Joiner, 2004; Rourke & Anderson, 2002; Salovaara, 2005). In sum, these studies acknowledged that collaborating can lead to information sharing, problem solving, and evaluation feedback, which are the result of interaction of the understandings of those who participate in the group activities. While accompanied with these benefits, CSCL inevitably poses some challenges. Kreijns, Krischner, and Jochems (2002) were among the first to identify that the design of CSCL environments do not completely fulfill expectations on supporting interactive and coordinated group learning. In fact they found a pitfall came from the assumption that collaborative interaction is taken for granted and that it will automatically happen in a CSCL environment.

Regarding the pitfall raised by Kreijns et al. (2002), we believe one approach to enhancing learners' collaborative behaviors in CSCL is to identify characteristics that differentiate low collaborating face-to-face (FTF) groups from high collaborating FTF groups and then determine if groups in CSCL also have these characteristics. If they do, and if those characteristics can be properly managed, groups in CSCL may enjoy fruitful experience in collaborating online. Two flaws most frequently mentioned in FTF groups are social loafing and free riding (see Williams & Karau, 1991; Albanese & Van Fleet, 1985). Actually social loafing and free riding essentially share a similar characteristic that an individual who is not providing the maximum effort is due to feeling of dispensability to the group (Kidwell & Bennett, 1993). Therefore, both of which reduce learners' willingness to make contributions for the group, as such it may handicap engendering of collaborative behaviors. DeSanctis and Gallupe (1987) also noted that while dealing with unstructured problems does not required members of the group to be in the same physical location, it is required for them to be aware of one another and to perceive themselves as being part of the group. An individual who feels himself as a member of a group may internalize the group's merit as part of his own preference,

which will help overcome problems of collective behaviors and barriers to resource exchange. For example, parents often make sacrifices for their children; however, they may not refer the actions to a sacrifice. In this regard, whether one considers oneself part of a group determines whether the welfare of the group is important to the individual, therefore a sense of group seems play an important role in explaining occurrence of collective behaviors.

In this regard, to explain engendering of collaborative behaviors in CSCL, this study referred the previous study by Gray (1989) who indicated a dynamic nature of collaboration, thus, more specifically, proposes a two-dimension process, namely perceptual, and supportive dimension. Considering past research (De Dreu, 2007) appreciated that effectiveness of a virtual group is significantly affected by conflicts or difference of opinions and shared motivations, this study consider task conflict and norms of cooperation critical in explaining collaborative behaviors. Therefore, we refer them to factors of perceptual dimension. Although the corresponding factors of perceptual dimension in this study are not uniquely related to CSCL settings, the nature of the CSCL settings has made them more salient and critical than in FTF environments. Indeed, the special nature of telecommunication used in the CSCL can affect the effectiveness of exchanging information and gaining consensus on information meaning (Daft, Lengel, & Trevino, 1987). In addition, a personality trait is a distinguishing characteristic, which is an individual's relatively consistent way of thinking, feeling, and behaving across situations. Besides, the relationship between personality and behavior has been extensively investigated and examined, which also draws our attention on this linkage.

Additionally, past research agree that diversity among group members can cause variations in their attitude, beliefs, and behaviors, which in turn affect overall performance (Shaw, 1981; Hackman, 1987). Shaw (1981) suggested that individual characteristics of group members are important factors related to group effectiveness. Specifically, Hackman (1987) indicated that personality is likely to contribute to work-team effectiveness. More recently, literature on CSCL confirmed that individual's personality influence learner's participation level (Ellis, 2001; Tien, 2000; Palloff & Pratt, 2001). These results all suggested that one's personality plays a critical role in FTF groups, therefore we expect it is influential in CSCL groups as well. Seeing that personality profile of extraversion/introversion has been widely agreed to be the first "Big Five" personality factor and which is particularly a valid predictor for tasks involving social interaction (McCrae & Costa, 1992), this study employed learners' personality profile of extraversion/introversion as individuals' characteristic and regarded it as the foremost determinant of collaborative behaviors. In line with the assumption, this study proposed that (a) learners' personality profile of extraversion/introversion would influence the perception of group context factors such as task conflict and norms of cooperation and (b) learners' perception of group context factors will guide their supportive behavior in CSCL setting.

## **Theoretical Foundations**

Collaborative learning is understood as a dynamic process of social interaction and participation. In this dynamic process, learners are first required to recognize their cognitive styles and abilities through observing capable others or a community in which the learning occurs. In fact, it is not a novel idea that collaborative learning refers to a dynamic process. In her early definition of collaboration, Gray (1989, p.5) described collaboration as "a process through which parties who see different aspects of a problem can constructively explore their differences and search for solutions that go beyond their own limited vision of what is possible." She believes the process of collaboration is seldom simple and straight-forward. In this regard, Gray (1989) thus proposed a three-phase process to describe the dynamic nature of collaboration. The first phase, which she calls the prenegotiation phase, is dedicated to address arriving at a shared definition of the problem. During the second phase, the parties identify the interests which brought them to the table, determine how they differ from the interests of others, set directions and establish shared goals. The final step of the collaborative process is the implementation phase during which parties gain the support from others. In line with the definition, both first and second phases mainly concentrate on carefully defining and, if need be, redefining the issues involved before moving on to solutions. And the last phase determines what support the participants may gain.

Intuitively, the objective of CSCL should not only to duplicate the features of a FTF environment, but rather to create a more effective learning environment to enhance the development of perception of belonging to the group. This study suggests that the dynamic process of CSCL can be divided into two dimensions, perceptual dimension and supportive dimension, which are rooted from Gray's (1989) definition of dynamic collaboration process. As shown in Table 1, perceptual dimension of CSCL covers learners' perception of norms of cooperation and task conflict, as



these aspects mainly cope with the interaction between members and the group to which they belong. Supportive dimension of CSCL focuses on the collaborative behaviors which are mutually beneficial to the group; thus it deals with learners' workload sharing and knowledge sharing. In particular, the former dimension is much important because a sense of group seems play an important role in leading later supportive behaviors (DeSanctis & Gallupe, 1987; Kidwell & Bennett, 1993; Rourke, 2000). Therefore, perceptual dimension is worthy for us to shed light on it.

Many studies have documented that individuals may exert less effort when working collectively without sense of belonging to the group (DeSanctis & Gallupe, 1987; Kidwell & Bennett, 1993; Rourke, 2000). Obviously, collaborative learning is not always beneficial in that students may encounter a set of problems such as non-contributing group members, unequal workload, and personal/social conflicts between group members (Becker & Dwyer, 1998). To overcome the unwilling result, group members collaborating online need to consciously strive to engage in group activities to create a perception of belonging to the group, therefore they can gain supports and benefits from CSCL.

*Table 1. A two-dimension of CSCL process*

Dimensions	Descriptions	Factors included
Perceptual	Measuring extent of collaborative norms in the group	• Norms of cooperation
	Assessing extent of task conflict in the group	• Task conflict
Supportive	Measuring extent of sharing resources or assisting those deficient in their work	• Workload sharing
	Examining learners' new ideas which gain through peer collaboration by interpersonal discourse	• Knowledge sharing

### **Perceptual dimension: norms of cooperation and task conflict**

In fact, group activities conducted by members of problem solving groups, mainly deal with two types of pressures in achieving quality solutions and high solution acceptance (Maier, 1967). On the one hand, there is pressure on each member to contribute unique, and possibly controversial, information to maximize the group's resources. On the other hand, members of the group tend to believe that a strong solution acceptance is best achieved through conformity of opinions. Consequently, individuals in a problem-solving group usually engage in activities such as dealing with disagreement in tasks and shared expectations that constrain and drive the action of group members. Thereby, engaging in these activities may shape learners' perception about the group to which they belong. This study adopts Maier's idea to refer group activities to dealing with task disagreement and shared expectations because the nature of CSCL also mainly focuses on the group work. In addition, shared expectations in this study are more specifically defined as norms of cooperation not only because they are essentially kind of norms, but also because cooperation is the major concern of this study. We also drew on Maier's (1967) description of the pressure aspects of a group for defining group context that group member may perceive. Learners' perception of group context variables thus includes perception of task conflict and perception of norms of cooperation across all individuals within a group. Norms are informal rules of conduct for behaviors that are considered important by most group members. According to Coleman (1990), a norm exists when the socially defined right to control an action is held not by the actor but by others. Thus, it represents a degree of consensus in the social system. Along with this perspective, Bettenhausen and Murnighan (1991) conclude that group norms tell group members how they are expected to behave. As a result, groups with shared norms of how to behave would waste less time in group meeting and use their resources better by avoiding duplication of work by group members.

The notion that CMC can accentuate, rather than attenuate, normative influence has received consistent empirical supports (Coleman, Paternite, & Sherman, 1999; Postmes, Spears, & Lea, 1998; Reicher, Spears, & Postmes, 1995). In fact, this finding can be derived from two theoretical perspectives, including social identity model of deindividuation effects (SIDE) and uncertainty reduction theory (URT). According to SIDE, interaction via a computer network can actually heighten group salience and, hence, conformity to a group norm because anonymity in CMC may induce a shift in focus from individual identity to collective identity (Reicher, Spears, & Postmes, 1995). The principal reason is that the decreased visibility of individuality can shift the emphasis away from concerns about others' individuality within the group, and towards the shared communalities and group concerns (Postmes, Spears, & Lea, 1998). Researchers have also suggested that anonymity and physical isolation in CMC not

only allows group members to separate posters from specific ideas (Jessup, Connolly, & Tansik, 1990) but also “serves to focus individuals’ attention on the task to glean the most possible information from the text-only medium” (Coleman, Paternite, & Sherman, 1999, p.61).

Another theoretical perspective that CMC can enhance conformity to induce group norms is due to URT (Berger & Calabrese, 1975). In line with its assumption, information exchange serves as an input that enables individuals to predict and explain the other’s behaviors (Berger, 1988). Apparently, this assertion to reduce uncertainty would be fundamentally unique in CMC interactions. In fact, it might be argued that many CMC settings offer conditions where uncertainty is high because of, for instance, anonymity or a lack of course of proper actions. In the absence of uncertainty, observation of others engaging in a behavior is likely to be associated with individuals’ own understanding of the appropriate mode of conduct. Further, if one does not perceive uncertainty, one is not likely to seek out normative information via active or interactive means (Berger & Calabrese, 1975).

In sum, CSCL is a cyberspace under which a situation of uncertainty develops, though it is usually free from constraints of regulation or boundaries. However, several studies have shown that anonymous CMC may serve as a vehicle for strong normative influence in groups (Coleman, Paternite, & Sherman, 1999; Jessup, Connolly, & Tansik, 1990; Postmes, Spears, & Lea, 1998). As such, a norm of cooperation allows group members to conduct their activities in supportive ways that are consistent with group expectations, because an absence of social cues provides a context in which individual differences are obscured. Being unable to perceive the self and the other as individuals may accentuate the unity of the group, and cause individuals to be perceived as group members rather than as unique individuals (Postmes, Spears, & Lea, 1998). Therefore, norms of cooperation give group members a common perspective that enables them to develop similar perceptions and interpretations of actions, leading a high level of supportive behaviors.

Another indicator used in perceptual dimension is task conflict. Conflict exists when inconsistent activities occur (Deutsch, 1973), which means disagreement between people or groups. Collaborative learning helps learners to discuss and finally comes up with an acceptable solution or consensus on given tasks. The consensus can be obtained by different ways, one of which is by discussing and reviewing disagreements with the task from each other. We focus on the task conflict is because which can increase group members’ tendency to scrutinize task issues and to engage in deep and deliberate processing of task-relevant information (Jehn, 1995). Therefore, in this study we defined task conflict as the different viewpoints and opinions about the task being performed.

It’s widely accepted that learning arises from the opportunities for the group members to explore multiple representations or perspectives on a specific task. However, it’s not that easy to have things happen naturally. Rourke (2000) remarks that “if students are to offer their tentative ideas to their peers, if they are to critique the ideas of their peers, and if they are to interpret others’ critiques as valuable rather than as personal affronts, certain conditions must exist. Students need to trust each other, feel a sense of warmth and belonging, and feel close to each other before they will engage willfully in collaboration and recognize the collaboration as a valuable experience.” In other words, a sense of group seems to be the first step for collaborative learning, and then the process of the deliberation for reaching shared understanding can move on.

In the context of CSCL, communication is one of the processes most affected by CMC and yet is central to the success of deliberation. Although the CSCL is designed to provide a communication environment that is similar to FTF communication, past research suggested that these two environments are clearly not the same (Joiner, 2004). A number of characteristics of CSCL have potentially benefited the process of collaboration. First, group members can reflect on previous arguments and reply with a thought-out response. Second, it provides the opportunity for group members to post opinions simultaneously. Finally, groups interacting via CMC have more equal participation among members than groups interacting FTF (Joiner, 2004). In this regard, learners in CSCL are more likely to achieve an agreement on the task through explicit and complete discussion on everyone’s ideas. Because such reflective behaviors (i.e., providing and receiving explanations) may help students clarify misunderstanding and enhance mutual comprehension (Chi & Bassock, 1989), therefore can help learners fill in gaps in their understanding.

In an enduring relationship, groups are more willing and more able to provide deliberation and reflection through repeated interactions (Coleman, 1988). Through repeated interactions, group members thereby become more similar regarding how to complete the task and develop similar understanding. In this regard, individuals are more likely to understand what others think and what problems embarrass their members; therefore they can provide pertinent

resources and assistance. The process of eliminating disagreement in cognition fosters learning and the development of new and sometimes highly creative insights, leading the group to become more effective (De Dreu & West, 2001; Jehn, 1995), thus may enact intensive supportive behaviors.

### **Supportive dimension: workload sharing and knowledge sharing**

While accomplishing the task at hand, group members are forced to look for free riding or social loafing because which significantly increase the other group members' burden of completing the task. If one member free rides, others may notice and, so as not to be victimized by the free rider, may also free ride. Obviously, if all members free ride, the group task will not be finished. In this sense, there should be workload sharing to avoid social loafing or free-riding in the collaborative groups (Albanese & Van Fleet, 1985). Moreover, according to Vygotsky, learning is seen as a social and cognitive process in which different perspectives are incorporated. The contribution of different understandings leads to a new, shared knowledge. Accordingly, researchers on collaboration propose that knowledge is created as it is shared, and the more it is shared, the more knowledge is created (Rogoff, 1995; Lave & Wenger, 1991). Seeing that past research have confirmed that workload sharing and knowledge sharing play important roles either in task accomplishment or in collaborative learning, this study, thereby regarded them as factors of supportive dimension.

Workload sharing is dividing work equally among group members. In this study the concept of workload sharing was similar to helping behavior, which was defined as "voluntary actions intended to help or benefit another individual or group of individuals" (Eisenberg & Mussen, 1989, p.3). Helping other group members, such as sharing resources or assisting those deficient in their work, benefits the immediate work group and the organization (Anderson & Williams, 1996; Borman & Motowidlo, 1993). As CSCL provides social connections for individuals to collaborate online, indeed it also stimulates the resources embedded within the social connections, such as one's experience, time and effort. Through repeated interactions, individuals are able to access and mobilize such resources on the task at hand. Van Dyne et al. (1995) confirmed that when a group has strong norms of cooperation, members expect mutual support from one another to enhance task completion. In other words, groups with norms of cooperation will encourage members to assist other group members.

In line with the past research that documented the significance of informal interaction, a learning environment should encourage active participation, interaction, and dialogue among learners to engage in a process of knowledge construction (Lin Hsiao, 1996). However, there are reduced informal interactions in virtual settings that would normally take place in a traditional collocated environment (Ellison, 1999). Thus, individuals who have experienced coworking with distant colleagues reported that this solitary environment engenders a feeling of remote isolation (Turban, McLean, & Wetherbe, 1996). The nature of solitary of a virtual setting in which CMC is used for collaboration adversely affect individuals' commitment toward their task and their members (Civin, 1999). Accordingly, if a CSCL setting provides learners strong feeling of solitary, learners may not commit to the group task and then prefer to work alone, which in turn leads to less willingness to share other members' workload.

CSCL system is centered upon the communications and interactions of individuals to generate specific domain knowledge that enable them to perform common functions and to learn from, contribute to, and collectively build upon that knowledge (Lee et al., 2003). While CSCL bring learners together virtually from all over the world, knowledge sharing among them has not lived up to expectation. Pfeffer and Sutton (1999) found that knowledge management in many organizations only emphasize on technology, particularly information technology. Dixon (2000) also pointed out that "build it and they will come" is the myth of knowledge sharing. Clearly, the biggest challenge in fostering CSCL system is the willingness to share knowledge with other members. By understanding what stimulates members of groups to share knowledge, it is possible for us to distinguish the mechanism that eases the sharing of knowledge among learners.

Knowledge sharing occurs in CSCL settings is significantly different from that occurring in conventional contexts in several ways. First, CMC allows a much larger number of individuals to be connected by informational linkages with nearly no more extra effort, which is quite different from in the case of FTF. For instance, the effort to email a message to all contacts in the address is only slightly more than the effort in sending the message to just only one recipient. This reduction in the effort needed to reach out to others increases the number of occasions for individuals to share various information or knowledge. Second, in order to better construct a new knowledge, learners are

required to be exposed on various perspectives elaborated by others. Different perspectives help individuals to glean information/data for the purposes of knowledge construction. Nonetheless, the evanescent nature of ephemeral messages means that it must be captured immediately before it loses. In this regard, a CSCL system that can automatically records learners' message content encourages the learners to resolve their own disputes by using CMC for knowledge sharing.

The sharing of knowledge is interwoven with the processes of interaction of different understandings. It is not limited to simple information exchange, but is related to the influence among members as a result of more frequent and in-depth interactions. This influence process is necessary for achieving mutual understanding between groups. For instance, individuals usually seek to either influence others into accepting these ideas or be influenced by others' ideas and attitudes. One way the influence is developed is likely through a shared expectation. Because individuals expect payback for contributions to an exchange, the perception of reciprocal benefits leads to mutual influence and success in future group exchanges (Cohen & Bradford, 1989). In this view, the perception of reciprocal benefits, specifically, is analogous to norms of cooperation because of their similar substance. Integrating the abovementioned and in line with the past study that documented that knowledge sharing behavior is likely to be influenced by contextual forces (Yoo & Torrey 2002), we then suspect it will be enhanced by the two indicators in perceptual dimension because of a sense of group (DeSanctis & Gallupe, 1987; Kidwell & Bennett, 1993; Rourke, 2000).

### **Group composition**

Research studies on factors affecting learners' outcome in CSCL indicated the importance of group composition (Adrianson, 2001). This study focuses on learners' personality profile as one way to distinguish different group compositions because past research acknowledges that members' diversity can cause variations in their attitude, beliefs, and behaviors, leading to affect overall performance (Shaw, 1981; Hackman, 1987). The greatest advance in personality research has been the emergence and broad acceptance of the "Big Five" model of personality (Digman, 1990; Hogan et al., 1996). Extraversion, widely agreed to be the first "Big Five" personality factor, appears to be a valid predictor for tasks involving social interaction (McCrae & Costa, 1992). Accordingly, we supposed that personality traits play an important role in individuals' social-communicative competence (Busato et al., 1999; Smolensky et al., 1990; O'Hair et al., 1995), because learning to behave adequately in virtual communicative settings requires the active interaction of a student with certain personality characteristics with other members.

In their study, Smolensky et al. (1990) suggest that individual's personality profile of extraversion/introversion has a positive influence on verbal interaction in CMC environments. This coexistence of extraversion/ introversion is particularly relevant to the way of learning in CSCL setting. This is because the unique components of the CSCL may assist introverted learners to express themselves more freely on the net than they feel able to in an offline relationship. For instance, studies have reported that extraverts tended to speak more often with longer comments and to be more talkative than introverts (Borman & Motowidlo, 1993; McCrae & Costa, 1992). However, Maldonado et al. (2001) evaluated messages on CMC and found that introverted subjects send messages with more information than they do in FTF settings. While CMC is different on a number of criteria, there is no reason to believe that frequent and open contact between learners in CSCL does not influenced by individuals' psychological profiles (Amichai-Hamburger & Furnham, 2007).

Conceptually, collaborating via the Internet inherently requires more intensive and active communication among one another and participation in the group activities. In this regard, personality profile of extraversion/introversion seems important. In fact, extraversion as a communication style contains two important components i.e., talkativeness and enthusiasm (Borman & Motowidlo, 1993; McCrae & Costa, 1992; Smolensky et al., 1990). Groups that have an extraverted communication style are more likely to generate more intensive and active interaction, because the number of learners answering actively to a post is supposed to be high. Thereby extraverts are more likely to dominate in CMC to impact on social influence (Smolensky et al., 1990). Seeing that many others actively participate in the group activates and individuals use the group to glean information on how to cope and behave, therefore individuals collectively engender group norms.

Unlike individuals in FTF setting where they glean this information by observing and communicating with others, individuals may not perceive both types of awareness without intensive interaction with one another in virtual groups. This is especially important because in a CSCL environment individuals' emphasis can only be placed on the

content of one's historical posts. Along with intensive interaction with one another to create a common understanding of the work being done, group members understand and incorporate each others' differences in opinions and viewpoints (Moreland & Myaskovsky, 2000). Thus, extravert groups are believed to have high level of awareness of who knows what in the group, and group members become more efficient at solving problems, coordinating, and allocating responsibilities for the task. This is likely to result in smoother task implementation and accomplishment, and subsequent reduction in disagreement or conflict about the task.

## Research Methods

A total of one hundred undergraduate students enrolling in a business organizational behavior course in Taiwan participated in this study. Participants were divided into 20 work groups, each consisting of 5 members. Groups engaged in academic work for course credit and were functioning as work teams with performance incentives. Two groups were dropped from the study later because of incomplete questionnaire data. The analyses we reported were based on a sample of 90 individuals within 18 groups. The course required a total of eighteen-week of instruction, lasting from mid-September 2005 to late January 2006. Upon the completion of group formation during Week 4, each group chosen a real-world business case from the case list and all participants completed an on-line personality assessment instrument. Each of the business case describes a unique real-world problem with no explicitly provided analytical strategy and no singly correct answers. During Week 7, each participant handed in an individual case report in which he (she) propose an analysis on the case and summarize written solution on its problem task. During Week 14, participants completed group dynamics instruments, including perceptions on task conflict, norms of corporation, and workload sharing. Groups were required to hand in a group term paper describing the written solution and analysis for the case problem during Week 15. One question asking the plausible solution on the business case problem was included in the final examination, which was held during Week 16.

A sample screen of the discussion board system is shown in Figure 1. It includes two main sub windows; the upper window is for learners to read course material, and the lower window is for learners to post any message they want to discuss or share with their members. To discourage interference from nonmembers, the system separated the intra-group spaces of different groups and hindered outsiders from intruding. This would further facilitate group members' feeling of belonging and provide opportunities for them to collectively reflect on their group's norms. In sum, on the one hand, participants used this inter-group space to discuss different opinions proposed by other group members and read course content shared by instructors. On the other hand, groups can get together to cooperate or compete with one another. Under the virtual group space, individuals were more likely to refer to other members' behavior.

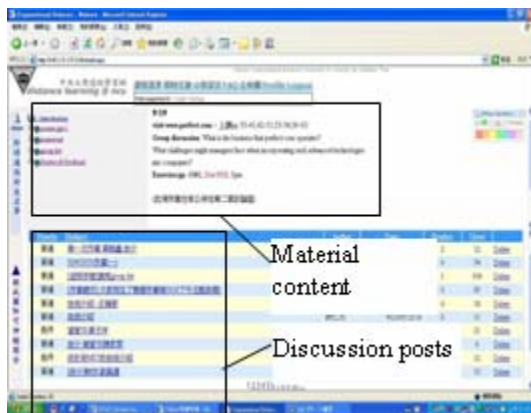


Figure 1. Sample screen of the discussion board

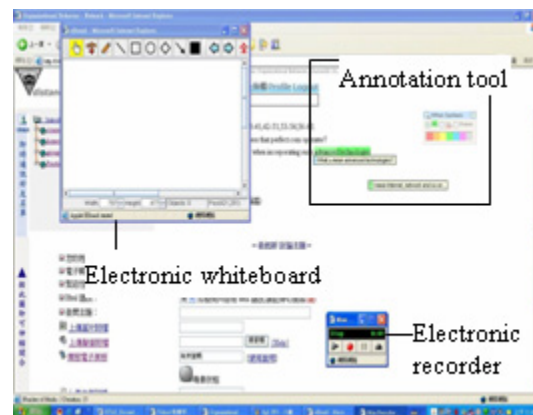


Figure 2. Available tools for posting and replying

In Figure 2, if participants click the button to post a message, the lower window will pop up the next interface asking them to fill in information such as subject, content and, if any, multimedia files. As shown in Figure 2, in the lower window, participants can chose to send a multimedia file which contains a drawing by using the electronic whiteboard or a voice by using the electronic recorder. The electronic whiteboard allowed all the group members to collectively come out a graphic metaphor illustrating the plausible model they developed. Because a computer-

generated visual aid could add interest to the group discussion, thus the electronic whiteboard enabled group members to have an effective brainstorming by drawing graphics asynchronously, and then made them more like a group. In addition, studies have pointed out that verbal cues are conducive to the collaborative or dialogic construction of knowledge in asynchronous communication context. The use of electronic recorder was then served as another supportive tool.

Annotation tool is another function that supports electronic brainstorming in our system. Participants used it to mark up any comment on the course or group work material and the system would share these comments to their group members. As such, when individuals move to the phase of evaluating others' ideas, they could easily understand what others think by referring to their members' annotations on the material, and then provided the opportunity for themselves to fill in gaps in disagreement with the task. Group members may be stimulated to build on the ideas of other group members. In sum, this system include one-to-one and one-to-many asynchronous text, image, audio, or video communications in its design.

## Measures

We use four measures to assess corresponding constructs. *Group Composition*. We used the extraversion/introversion subscale of Myers-Briggs Type Indicator (MBTI) instrument (Myers, 1984) to measure subject's extravert and introvert type. Participants were asked to judge the accuracy level of each adjective as a description of self from one of two options. A person's MBTI score determines his or her type, a label based on his or her dominant preference for the extraversion/introversion dimension. Accordingly, three team composition types were classified based on the proportion of extraverts in the team: introvert team (0-1 of extroverts), extravert team (4-5 of extraverts), and hybrid team (2-3 of extraverts).

*Task Conflict*. Task conflict was measured by 3 items adapted from Jehn (1995) and Jehn and Shah (1997). The items were assessed on a 5-point Likert scale. A sample item of task conflict is "To what extent are there differences of opinion in GROUP?" *Norms of cooperation*. We measured norms of cooperation with three items adapted from Wageman's (1995) cooperation norm scale. A sample item is "In my group, we think that everyone should volunteer to do things for the group". *Workload sharing*. Individual's workload sharing was referred to and measured by Van Dyne and LePine's (1998) scale. The scale includes 4 items. A sample item is "We regularly take time to figure out ways to improve our group's work processes". *Knowledge Sharing*. In order to measure group knowledge sharing, we operationalized two variables: individuals' pre-discussion knowledge and individuals' post-discussion knowledge. The former is counting the number of feasible solutions proposed in individual's case report. The latter we included a question asking participants to list again all the feasible solutions in the final examination at the end of the semester. The individual gained one point of knowledge sharing if a solution which hadn't revealed in one's own pre-discussion knowledge but had appeared in other members' pre-discussion knowledge was presented in one's final examination. Finally we accumulated each group members' knowledge sharing points to serve as the indicator of group knowledge sharing.

## Results and Discussion

The reliability coefficients of task conflict, norms of cooperation, and workload sharing, measured by Cronbach's  $\alpha$ , were 0.813, 0.801 and 0.837 respectively. In this study three types of group composition were classified: hybrid group, introvert group, and extravert group. Table 2 presents the descriptive statistics associated with our research variables for each type of group.

Table 2. Descriptive statistic of the study variables

Dimensions of CSCL Process	Perceptual		Supportive	
	Norms of cooperation	Task Conflict	Workload sharing	Knowledge sharing
Group type (# of group)	Mean (std)	Mean (std)	Mean (std)	Mean (std)
Hybrid groups (n=10)	4.13 (0.29)	3.04 (0.25)	4.01 (0.31)	3.55 (2.42)
Introvert groups (n=5)	4.13 (0.18)	2.87 (0.29)	4.09 (0.25)	2.58 (1.58)
Extravert groups (n=3)	4.60 (0.12)	2.50 (0.22)	4.48 (0.20)	3.00 (3.50)

Of the three types of group, hybrid groups had the highest task conflict, the lowest workload sharing, and moderate group performance. Introvert groups had moderate task conflict and workload sharing. Extravert groups had the lowest group task conflict and the highest level of workload sharing and cooperative norms. The correlation matrix shown in Table 3 indicates cooperative norms was positively associated with workload sharing but was negatively associated with task conflict. Workload sharing was negatively associated with task conflict. None of the relationship between knowledge sharing and other variables was significant in this study.

Table 3. Correlation matrix of study variables

	1	2	3	4
1. Norms of cooperation	1.00			
2. Task conflict	-0.58*	1.00		
3. Workload sharing	0.73**	-0.69**	1.00	
4. Knowledge sharing	0.25	0.13	0.09	1.00

\*\*·: Correlation is significant at the 0.01 level; \*·: Correlation is significant at the 0.05 level

### The impact of group composition on perceptual dimension

One-way ANOVA test was conducted to examine the effect of team composition on learners' perceptual dimensions of CSCL process, namely task conflict and norms of cooperation. Table 4 shows significant group differences on task conflicts ( $F = 4.92, p < 0.05$ ). Turkey's LSD post hoc analysis was employed to further identify which groups differ significantly. As a result, Table 5 shows that the difference on task conflict between extravert group and hybrid group is statistically significant ( $p < 0.05$ ). Extravert groups had the lowest task conflict whereas hybrid groups perceived the highest task conflict. Regarding the group norms of cooperation, Turkey's LSD post hoc analysis indicated that extravert groups perceived a significantly higher norms of cooperation than both hybrid and introvert groups.

Table 4. Effects of team composition on learners' perceptual dimension of CSCL process

Dependent Variable	Source	SS	df	MS	F	Sig.
Task conflict	Between Groups	0.66	2	0.33	4.92	0.02*
	Within Groups	1.01	15	0.07		
	Total	1.67	17			
Norms of cooperation	Between Groups	0.55	2	0.27	4.51	0.03*
	Within Groups	0.91	15	0.06		
	Total	1.46	17			

\*·:  $p < 0.05$

Table 5. Multiple comparisons on norms of cooperation and task conflict

Dependent Variable	(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.
Norms of cooperation	H	I	0.00	0.14	0.98
		E	-0.47	0.16	0.01*
	I	E	-0.47	0.18	0.02*
Task conflict	H	I	0.17	0.36	0.06
		E	0.53	0.43	0.01*
	I	E	0.36	0.48	0.43

\*·:  $p < 0.05$

“H” stands for hybrid groups; “I” stands for introvert groups; “E” stands for extravert groups

### Relationship between learners' perceptual dimension and supportive dimension

Regression analysis was used to investigate whether learners' perceptual dimension of CSCL process would lead to supportive behaviors, namely workload sharing and knowledge sharing. Table 6 indicates that the effects of norms of cooperation and task conflict on workload sharing were significant ( $\beta = 0.48, p < 0.05$ ;  $\beta = -0.40, p < 0.05$ ,

respectively). This result indicated that groups perceiving stronger norms of cooperation and less task conflict resulted in more workload sharing. However, neither norms of cooperation and task conflict were significant in explaining knowledge sharing. Accordingly learners' collaboration process of perceptual dimension can only partially predict one of the indicators in the supportive dimension, namely workload sharing.

*Table 6. Regression analysis of variables of perceptual dimension on workload sharing*

Independent Variable	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Norms of cooperation	0.53	0.21	0.48	2.58	0.02*
Task conflict	-0.41	0.19	-0.40	-2.15	0.04*

Dependent Variable: Workload sharing

$R^2 = 0.63$ ,  $\text{Adj-}R^2 = 0.58$

\*:  $p < 0.05$

The present study has two aims: (1) to test whether learners' personality profile of extraversion/introversion would influence learners' their perception of group context variables such as task conflict and norms of cooperation and (2) to test if these variables predict learners' supportive behaviors in the CSCL setting. After testing the two questions, we believe our study makes several contributions. First, we propose a two-dimension process, perceptual dimension and supportive dimension, to explain engendering of collaborative behaviors in CSCL. Second, we demonstrate the importance of group composition and show its impact on learners' perceptual dimension. Finally, rather than identifying the relationship between perceptual and supportive dimension, this study corroborate there exist a perception of group context that affects learners' willingness to conduct collective behaviors in CSCL.

This study explored and identified two possible factors affecting learners' supportive behaviors and then integrated Gray's (1989) definition of collaboration to propose a two-dimension process for explaining how collaborative behaviors were engendered. The results reveal that learners' supportive behaviors (workload sharing) can be enhanced by perception of group context (norms of cooperation and task conflict). Based on a wide review of identifying critical antecedents leading to collaborative behaviors in CSCL, this study suggests perceptual dimension as the first stage of enhancing collaboration in CSCL. After experiencing high cooperation and low task conflict, learners perceive a sense of group and are more willing to contribute their effort to share workloads.

Extraversion contains an element of positive affectivity (George, 1992), which is an overall sense of well-being. Accordingly, an extravert group develops norms of cooperation that are likely to persist as the group moves toward task completion. Furthermore, a low level of task conflict in extravert groups indicated that different viewpoints on a specific task seldom emerged. Compared with hybrid groups, the result was consistent with the suggestion from previous literature that increasing group diversity can lead to conflict between members (Jehn 1995). Task conflict exists when group members disagree about the content of the tasks, aid differ in viewpoints, ideas, and opinions. This suggests that learners in homogeneous group composition are more likely to reach a shared cognition, which can reduce disagreement.

We propose a two-dimension process and test the link between them. In fact, by testing the relationship, we documented that learners' perception of group context variables play an intermediate role in controlling their later collaborative behaviors. It not only help us explain how collective behaviors were engendered, but implies a shift in focus from duplicating parallel functions in FTF setting to maintaining social and psychological perception of group context as designing a CSCL system. Along with the consecutive advance in information technology, developers can now provide more novel and parallel mechanisms to duplicate features in FTF. While accepting its usefulness in fulfilling learners' requirement to study online, we suggest educators and developers should not restrict learning mechanisms to instructional functions aimed at cognitive processes. In contrast, perceptual functions aimed at social and psychological processes should also be emphasized and established. We can't tell which variables other than this study related to group context are required to be facilitated and monitored by information technology in that only two indicators of perceptual dimension were used in this study. However, it is reasonable to suppose that other critical variables may shape learners' perception of group context. If we can identify more of these factors and determine the predictive power of each in explaining learners' supportive behaviors, it is possible for instructors and educators to manage and enhance behaviors conducive to online collaborative learning.



## Conclusion

This paper proposes a two-dimension of dynamic process for further our understanding of what factors lead learners to conduct collective behaviors of collaboration in CSCL settings. From the empirical evidence, the impact of group composition of extraverts/introverts on learners' perception of task conflict and perception of norms of cooperation has been supported. It was also supported that learners experiencing high level of shared norms of cooperation and low level of task conflict are more likely to conduct supportive behaviors of workload sharing.

Our results indicated that extravert groups are more likely to reach shared norms of cooperation in CSCL setting. Extraversion usually affects interpersonal relations through the quality of social interactions. As such, extraverts are active participants in group interactions, even in CSCL setting; therefore it seems that they play an important role in explaining engendering of collaborative behaviors. In this regard, we believe extraverts may have high intragroup popularity to influence others to accept their ideas and are more likely to perceive normative information and to match other members' expectations. Further, there is another alternative. Given that extraverts are active and usually adept at interacting, it is reasonable to suspect that they are more likely to establish norms of cooperation rather than to follow it. Our study cannot rule out this alternative, future research may take into account the influence of norms establishers and norms followers while considering questions about collective behaviors.

Our study found that learners' supportive behaviors of workload sharing were stimulated by enhancing their perception of group context variables. The coefficient of norms of cooperation indicated that learners are more willingness to share resources or assist those members deficient in their work as they perceived strong norms of cooperation. As one begin to interact and coordinate to finish the group task over time, perceiving cooperative norms would trigger one's supportive actions to help others. As a result, shared norms of cooperation lead to the positive attitude of workload sharing. Task conflict was proposed as a key source of divergent thinking, encouraging use of unique information, and pooling of resources to create a better solution. However, this study found that it is negatively associated with group members' supportive behaviors. Contrary to the effect of norms of cooperation, this result suggests that learners' perceptions of group context not always positively affect their supportive behaviors. Collaborative learning provides equal power relations among peers allow learners to actively take different perspectives (Matusov & Hayes, 2000), leading them to review their own ideas and coordinate actions and perspectives to resolve cognitive disagreements. If these cognitive differences can not be resolved, learners in the group may perceive a high level of task conflict, which in turn decrease willingness to share workload of other members.

Knowledge sharing was not significantly associated with norms of cooperation and task conflict in this study. This unexpected result brings us a question why the relationship was not supported. Compared with prior literatures that confirmed knowledge sharing comes from interactions and clarifications (Lee et al., 2003; Dixon, 2000), our result seems to be contradictive. Apparently there might be other factors leading to intensive knowledge sharing. One of the plausible factors may be derived from the well-known motivation theory. Maslow (1954) suggests that human conduct is motivated by five classes of needs: basic, safety, belongingness, esteem, and self-actualization. Davenport and his colleagues (1998) also argue individuals will not share their knowledge as they think their knowledge is valuable and important, unless a sound incentive system is available. Thus, future research may consider integrating motivational factors into the two-dimension dynamic process. Another plausible explanation stems from media richness theory (Daft et al., 1987), which argues that media are differentially suited to diverse tasks. Asynchronous technologies, such as discussion forum, are useful for arranging meetings and sharing provisional documents, while synchronous technologies are more useful for brainstorming and decision-making (Rourke & Anderson, 2002). The task employed in this study is to solve a real-world business case required a brainstorming from group members to address its unstructured problems. Asynchronous technologies seem to be lack of establishment of brainstorming or be less efficient in this way. As a result, knowledge sharing was not significant as what we had expected. Future research should take Rourke and Anderson's suggestion into account and try to provide synchronous technologies to see whether learners with synchronous technologies support have more knowledge sharing behaviors.

## Implications

From a practical standpoint, our results provide good news for instructors and educators, because in order to stimulate supportive behaviors in CSCL they can affect learners' perception of group context variables (norms of

cooperation and task conflict) by managing group composition. We use personality profiles of extraversion/introversion as group composition, which means instructors and educators may wish to assign more extravert learners to the group because they can stimulate the enactment of norms of cooperation and help reach an agreement among members on the task at hand. It has been widely emphasized that CSCL brings many advantages for students such as flexibility in terms of time and space. However, it is also a double-edged sword in educational domain. Although it may accompany with a number of advantages, it also hinder teachers to identify each learners' characters by extending psychological distance as well. Through this study, learners' impact of psychological distinction on attitude toward the cyber group has been supported, highlighting its indirect influence on collective behaviors online. In this regard, educators and instructors should be aware of learners' uniqueness behind the fascinating functions enhanced by information technology.

Finally, learners in homogeneous group composition are more likely to reduce disagreement of how to complete the task and increase goodwill in supporting members because of a shared agreement of what should be done. Although not significantly supported, heterogeneous groups seem perform better knowledge sharing. It may imply a contingency for educators and instructors to determine a suitable group composition while considering the task characteristics. For instance, a routine task may not require too much information exchange, but may rely on its members to share required resources. In addition, a non-routine task may require its members to exchange valuable ideas to solve the unstructured problems properly. In sum, we suggest that although CSCL definitely brings learners flexibility in learning, it doesn't bring educators stability in teaching.

### **Limitations of the study**

The main limitation in this study was the insufficient sample size. We have a total of 100 subjects (in 20 groups) participating in the study. However, only the data of 90 subjects (in 18 groups) were employed for further analysis due to data missing. Accordingly, the generalizability of our study results is limited. In addition, our study focused on the non-routine, problem-solving based task. This specific task characteristic could become our study limitation. Future study should employ other types of tasks to examine if different types of task would result in different stories. Finally, many other interesting factors are excluded in this study because of the limitation of research scope. As a result, future research may consider effects of some other factors on individuals' perception of group context and the latter influence on collaborative behaviors. One of the factors excluded from this study, for instance, is ethnicity. Different ethnicity has different culture that provides basis of how we assert, question, demand, and judge. In fact, a culture is a complex set of shared beliefs, values, and concepts which enables a group to make sense of its life and which provides it with directions for how to behave. Similarly, although not so complicated as the issue of ethnicity, individuals' gender and age are acknowledged to be critical antecedents in shaping ones' courses of action. Therefore, a follow-up study dealing with issues of ethnicity, gender, and age seems to be worthwhile.

### **Reference**

- Adrianson, L. (2001). Gender and computer-mediated communication: Group processes in problem solving. *Computers in Human Behavior*, 17 (1), 71-94.
- Albanese, R., & Van Fleet, D. D. (1985). Rational behavior in groups: The free-riding tendency. *Academy of Management Review*, 10 (2), 244-255.
- Altinay, L., & Paraskevas, A. (2007). A computer-supported collaborative learning (CSCL) approach in teaching research methods. *International Journal of Hospitality Management*, 26 (3), 623-644.
- Amichai-Hamburger, Y., & Furnham, A. (2007). The positive net. *Computers in Human Behavior*, 23 (2), 1033-1045.
- Anderson, S. E., & Williams, L. J. (1996). Interpersonal, job, and individual factors related to helping processes at work. *Journal of Applied Psychology*, 81 (3), 282-296.
- Becker, D., & Dwyer, M. (1998). The impact of student verbal/visual learning style preference on implementing groupware in the classroom. *Journal of Asynchronous Learning Networks*, 2 (2), 61-69.

- Berger, C. R. (1988). Uncertainty and information exchange in developing relationships. In S. W. Duck (Ed.), *Handbook of personal relationships: Theory, research and interventions* (pp. 239-255), Chichester: Wiley.
- Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1 (2), 99-112.
- Bettenhausen, K. L., & Murnighan, J. K. (1991). The development of an intragroup norm and the effects of interpersonal and structural challenges. *Administrative Science Quarterly*, 36 (1), 20-35.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection* (pp. 71-98), San Francisco: Jossey-Bass.
- Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (1999). The relation between learning styles, the Big Five personality traits and achievement motivation in higher education. *Personality and Individual Differences*, 26, 129-140.
- Chi, M. T. H., & Bassock, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251-282), Hillsdale, NJ: Erlbaum.
- Civin, M. A. (1999). On the vicissitudes of cyberspace as potential space. *Human Relations*, 52 (4), 485-506.
- Cohen, A. R., & Bradford, D. L. (1989). Influence without authority: The use of alliances, reciprocity, and exchange to accomplish work. *Organizational Dynamics*, 17 (3), 4-18.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95-120.
- Coleman, J. S. (1990). *Foundations of social theory*, Cambridge: Harvard University Press.
- Coleman, L. H., Paternite, C. E., & Sherman, R. C. (1999). A reexamination of deindividuation in synchronous computer-mediated communication. *Computers in Human Behavior*, 15 (1), 51-65.
- Daft, R. L., Lengel, R. H., & Trevino, L. (1987). Message equivocality, media selection, and manager performance. *MIS Quarterly*, 11 (3), 355-366.
- Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organization manage what they know*, Boston, MA: Harvard Business School Press.
- De Dreu, C. K. W. (2007). Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *Journal of Applied Psychology*, 92 (3), 628-638.
- De Dreu, C. K. W., & West, M. A. (2001). Minority dissent and team innovation: The importance of participation in decision making. *Journal of Applied Psychology*, 86 (6), 1191-1201.
- DeSanctis, G., & Gallupe, R. B. (1987). A foundation for the study of group decision support systems. *Management Science*, 33 (5), 589-609.
- Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*, New Haven, CT: Yale University Press.
- Digman, J. M. (1990). Personality structure: An emergence of the five-factor model. *The Annual Review of Psychology*, 41 (1), 417-440.
- Dixon, N. M. (2000). *Common knowledge: How companies thrive by sharing what they know*, Boston, MA: Harvard Business School Press.
- Eisenberg, N., & Mussen, H. M. (1989). *The roots of prosocial behavior in children*, Cambridge: Cambridge University Press.
- Ellis, A. (2001). Student-centred collaborative learning via face-to-face and asynchronous online communication: What's the difference? Paper presented at the Meeting of the Eighteenth Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, December 9-12, 2001, Melbourne, Australia.
- Ellison, N. B. (1999). Social impacts: New perspectives on telework. *Social Science Computer Review*, 17 (3), 338-356.
- George, J. M. (1992). The role of personality in organizational life: Issues and evidence. *Journal of Management*, 18 (2), 185-213.

- Gray, B. (1989). *Collaborating: Finding common ground for multiparty problems*, San Francisco: Jossey-Bass Inc.
- Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315-342), Englewood Cliffs, NJ: Prentice-Hall.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, 51 (5), 469-477.
- Jehn, K. A. (1995). Multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 40, 256-282.
- Jehn, K. A., & Shah, P. P. (1997). Interpersonal relationships and task performance: An examination of mediating processes in friendship and acquaintance groups. *Journal of Personality and Social Psychology*, 72 (4), 775-790.
- Jessup, L. M., Connolly, T., & Tansik, D. A. (1990). Toward a theory of automated group work: The deindividuating effects of anonymity. *Small Group Research*, 21 (3), 333-348.
- Joiner, R. (2004). Supporting collaboration in virtual learning environments. *CyberPsychology & Behaviour*, 7 (2), 197-200.
- Kidwell, R. E., & Bennett, N. (1993). Employee propensity to withhold effort: A conceptual model to intersect three avenues of research. *Academy of Management Review*, 18 (3), 429-456.
- Kreijns, K., Kirschner, P., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society*, 5 (1), 8-22.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*, Cambridge: Cambridge University Press.
- Lee, F., Vogel, D., & Limayem, M. (2003). Virtual community informatics: A review and research agenda. *Journal of Information Technology Theory and Application*, 5 (1), 47-61.
- Lin Hsiao, J. W. D (1996). *CSCL theories*, Retrieved Aug 5, 2008, from <http://www.edb.utexas.edu/csclstudent/Dhsiao/theories.html>.
- Maier, N. R. F. (1967). Assets and liabilities in group problem-solving: The need for an integrative function. *Psychological Review*, 74 (4), 239-249.
- Maldonado, G. J., Mora, M., Garcia, S., & Edipo, P. (2001). Personality, sex and computer communication mediated through the Internet. *Anuario-de-Psicologia*, 32, 51-62.
- Maslow, A. (1954). *Motivation and personality*, New York: Harper.
- Matusov, E., & Hayes, R. (2000). Sociocultural critique of Piaget and Vygotsky. *New Ideas in Psychology*, 18 (2-3), 215-239.
- McCrae, R. R., & Costa, P. T. (1992). A five-factor theory of personality. In L. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp.139-153), New York: Guilford.
- Moreland, R. L., & Myaskovsky, L. (2000). Explaining the performance benefits of group training: Transactive memory or improved communication? *Organizational Behavior and Human Decision Processes*, 82 (1), 117-033.
- Myers, I. B. (1984). *Myers Briggs type indicator*, Palo Alto Ca: Consulting Psychologists Press.
- O'Hair, D., Friedrich, G. W., Wiemann, J. M., & Wiemann, M. O. (1995). *Competent communication*, New York: St. Martin's Press.
- Palloff, R. M., & Pratt, K. (2001). *Lessons from the cyberspace classroom: The realities of online teaching*, San Francisco: Jossey-Bass.
- Pfeffer, J., & Sutton, R. (1999). Knowledge "What" to do is not enough: Turning knowledge into action. *California Management Review*, 42 (1), 83-108.
- Postmes, T., Spears, R., & Lea, M. (1998). Breaching or building social boundaries? SIDE-effects of computer-mediated communication. *Communication Research*, 25 (6), 689-715.

- Reicher, S. D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (pp. 161-198), Chichester, UK: Wiley.
- Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J. V. Wertsch, P. del Rio & A. Alvarez (Eds.), *Sociocultural studies of mind* (pp.139-163), Cambridge: Cambridge University Press.
- Rourke, L. (2000). Operationalizing social interaction in computer conferencing. *Paper presented at the 16th Annual Conference of the Canadian Association for Distance Education*, May 3-6, 2000, Quebec City, Canada.
- Rourke, L., & Anderson, T. (2002). Using web-based, group communication systems to support case study learning at a distance. *The International Review of Research in Open and Distance Learning*, 3 (2), 1-13.
- Salovaara, H. (2005). An exploration of students' strategy use in inquiry-based computer supported collaborative learning. *Journal of Computer Assisted Learning*, 21 (1), 39-52.
- Shaw, M. E. (1981). *Group dynamics: The psychology of small group behavior* (3<sup>rd</sup> Ed.), New York: McGraw-Hill.
- Smolensky, M. W., Carmody, M. A., & Halcomb, C. G. (1990). The influence of task type, group structure and extraversion on uninhibited speech in computer-mediated communication. *Computers in Human Behavior*, 6 (3), 261-272.
- Tiene, D (2000). Online discussions: A survey of advantages and disadvantages compared to face-to-face discussions. *Journal of Educational Multimedia and Hypermedia*, 9 (4), 369-382.
- Turban, E., McLean, E., & Wetherbe, J. (1996). *Information technology for management: Improving quality and productivity*, New York: Wiley.
- Van Dyne, L., Cummings, L. L., & McLean Parks, J. (1995). Extra-role behaviors: In pursuit of construct and definitional clarity. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 215-285), Greenwich, CT: JAI.
- Van Dyne, L., & LePine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal*, 41 (1), 108-119.
- Wageman, R. (1995). Interdependence and group effectiveness. *Administrative Science Quarterly*, 40, 145-180.
- Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology*, 61 (4), 570-581.
- Yoo, Y., & Torrey, B. (2002). National culture and knowledge management in a global learning organization. In C. W. Choo & N. Bontis (Eds.), *The strategic management of intellectual capital and organizational knowledge* (pp. 421-434), Oxford, England: Oxford University Press.

## Effects of Practice in a Linear and Non-linear Web-based Learning Environment

Florence Martin

University of North Carolina, Wilmington, NC, USA // florencemartin@gmail.com // Tel: +1 (480) 650-6926

### ABSTRACT

Instructional elements remain the foundation of current instructional design practice. Practice is the instructional element provided after learners have been given information required to master an objective and Navigation is the non-instructional element guides the learner on the sequence of instruction. Linear Navigation can be referred to as program control where the learners do not have control over sequence and Non-Linear Navigation can be referred to as Learner control over sequence of instruction. The purpose of this study was to investigate the effect of practice with feedback, navigation type on achievement, attitude, and time when students use a web-based instructional program and the interaction between practice and navigation type. 240 students from a large south western university participated in four different web based environments. Significant differences were found for practice main effect, but not for navigation main effect. There was no interaction between practice and navigation. There were significant differences for attitude items, I learned a lot from this program and the program gave me enough opportunity to practice between the treatments who received practice and those who did not. This study reinforces on the importance of practice and has implications for the design and development of web-based, multimedia instruction.

### Keywords

Practice, Navigation, Web-based instruction, Linear and non-linear, Instructional elements

### Introduction

#### Instructional elements

Forty years ago, Robert Gagné published the first edition of his book *The Conditions of Learning* (1965) in which he proposed nine events of instruction that provide a sequence for organizing a lesson. These events remain the foundation of current instructional design practice (Reiser, 2002; Richey, 2000). They represent desirable conditions in an instructional program and increase the probability of successful learner achievement (Gagné, 1965, 1985, 1988; Gagné, Briggs & Wager, 1992). Other authors cite similar elements of instruction that promote student learning from an instructional program (Dick, Carey, & Carey, 2005; Sullivan & Higgins, 1983).

According to Forcier & Descy (2002), “every learning environment has an implied method of information presentation” (p. 104). Information is necessary to perform the task stated in an objective and is presented in a straightforward manner (Sullivan & Higgins, 1983). Apart from the information which is the basic instructional element that is needed in any instructional program, other instructional elements such as objectives, examples, practice with feedback and review are provided in addition. Some of the instructional elements that have been suggested by Gagne (1985) and Dick et al. (2005) to promote learning are objectives, practice with feedback, examples and review.

According to Clark & Mayer (2007) instructional methods are “the elements included in instruction for the purpose of supporting the achievement of the learning objective . . . instructional methods are intended to encourage learners to use appropriate cognitive processing during instruction” (p. 314). These authors indicate that multimedia will promote learning to the extent that it supports human cognitive processes. Each of these instructional elements is described in the following paragraphs.

An instructional *objective* is a statement that describes an intended outcome of instruction (Dick et al., 2005; Mager, 1997). Objectives facilitate cognitive processing by focusing student attention, directing selective perception of specific lesson content, communicating expectations, and organizing new information into an existing structure (Foshay, Silber, & Stelnicki, 2003; Gagné, 1985; Gagné, Wager, Golas, & Keller, 2005, Smith & Ragan, 2005). According to Reiser & Dick (1996), “At a fairly early stage, learners should be informed of what it is that they are going to be able to do when they finish the instructional process. By knowing what will be expected of them, learners may be better able to guide themselves through that process” (p.48). Morrison, Ross, & Kemp (2006), indicate that

although the general trend continues to be the use of objectives as a pre-instructional strategy, research results suggest providing learners with objectives is not as effective as once thought.

*Practice* involves eliciting performance from learners (Gagné, 1985; Gagné et al., 2005). It is often provided after learners have been given information required to master an objective. Practice provides an opportunity for learners to strengthen new knowledge by internalizing it so they can recall and use it (Foshay et al., 2003). It helps to confirm correct understanding and repeated practice increases the likelihood of retention (Klein, Spector, Grabowski & de la Teja, 2004; Kruse & Kevin, 1999). Practice is effective when it is aligned with assessment and with the skills, knowledge and attitudes reflected in the objectives (Merrill, 2002; Reiser & Dick, 1996).

*Feedback* can be defined as “knowledge of one’s performance provided” (Delgado & Prieto, 2003, p. 73). Practice provides an opportunity for feedback that confirms the student’s answer as being correct or indicates that it is incorrect. Feedback strengthens the probability of correct responses and reduces the probability of subsequent incorrect responses (Philips, Hannafin & Tripp, 1988). Kulhavy & Stock (1989) define feedback as information consisting of two components: verification and elaboration. Verification is the simple, dichotomous judgment that an initial response was right or wrong. Elaboration consists of all substantive information contained in a feedback message.

*Examples* are verbal or graphical information that provides additional clarification of rules or information presented to learners. Kruse & Kevin (1999) include examples, non-examples, graphical representation and analogies as guidance strategies that can be used to further clarify new content that is presented.

*Review* typically provides an outline of the key information that was presented to learners. It is intended to reinforce learning, at the end of the instruction, often just before students are tested. Reiser & Dick (1996) cite the value of reviews to bring closure to instruction and to help reinforce the skills and knowledge students should have acquired. Mattiske (2001) suggests that a review activity immediately after participants have learned something new reassures them that they are learning. Klein et al. (2004) suggest that learners should be given time to reflect and review after new information has been presented to them. Gagné et al., (2005) indicate that spaced reviews should be given to learners to help them retrieve and use newly acquired information.

## **Practice and Feedback**

Researchers have found that practice has a significant effect on performance. Hannafin (1987) reported a significant difference between practiced and non-practiced items on the learning of cued and uncued information presented via computer-based instruction. Phillips et al. (1988) found a significant difference favoring practice over no practice in an interactive video in which practice items were embedded questions. Hannafin, Philips, Rieber & Garhart (1987) noted that practice effects were more pronounced for facts than for applications in computer-based instruction. Participants who received intellectual skills practice in a cooperative learning environment performed significantly better than those who received verbal information practice (Klein & Pridemore, 1994).

Practice provides an opportunity for feedback that confirms the student’s answer as being correct or indicates that it is incorrect. This feedback strengthens the probability of correct responses and reduces the probability of subsequent incorrect responses. (Philips et al., 1988). Simple forms of feedback are effective when learners are able to answer items correctly. More elaborate forms such as providing and explaining the correct answer and explaining why a wrong answer is incorrect are helpful when learners answer incorrectly (Kulhavy, 1977). Simple forms of feedback are most effective for simple verbatim and verbal information types of learning (Kulhavy, White, Topp, Chan & Adams, 1985).

## **Non-Instructional elements**

Apart from the instructional elements which Gagne (1985), Dick et al. (2005) have proposed and many researchers have researched on, there are the non-instructional elements such as usability, navigation type, learner control that also influence learning. Researchers have examined the effects of these non-instructional elements such as navigation

type (Su & Klein, 2006), personalization (Ku & Sullivan, 2002), animated agents (Atkinson, Mayer & Merrill, 2005) in student achievement and attitudes.

As Hannafin (1987) noted, some design strategies may have positive effects when used in isolation that are diminished or negated when these strategies are used in combination with more powerful techniques. The effects of these instructional elements could be enhanced or diminished when used in combination with other variables such as different navigation types, media types.

### **Linear and Non-Linear Navigation (Program Control versus Learner Control over the sequence of instruction)**

Linear Navigation can be referred to as program control where the learners do not have control over sequence and Non-Linear Navigation can be referred to as Learner control over sequence of instruction. Hypertext has been defined as an approach to information management in which data is stored in a network of nodes connected by links. Shneiderman defines hypertext as "a database that has active cross-references and allows the reader to "jump" to other parts of the database as desired" (Shneiderman & Kearsley, 1989). Much of the previous research on the effects of navigation tools look at efficiency and effectiveness of hypertext environment (Boechler & Dawson, 2002; DeLucas & Larkin, 1995; Dias & Sousa, 1997; McDonald & Stevenson, 1998). Efficiency measures are based on speed and the number of steps taken to complete an information search. Effectiveness measures focus on the user's search accuracy and the users understanding of the structure of the document (Boechler & Dawson, 2002). Different types of navigation structures are available – hypertext links, hypertext links with graphics, keyword index (Hammond & Allison, 1989). Studies were also conducted to test for the most effective navigation types such as content list and content map (Su & Klein, 2006).

Many of the instructional programs designed to test in computer based instruction were built with Hypercard and Tool book initially and now it is built with Director, Authorware, Dreamweaver and Flash. In general these programs have been linear in format (Freitag & Sullivan, 1995; Schankenberg, Sullivan, Leader & Jones, 1995; Martin, Klein & Sullivan, 2006). These programs do not allow learners to navigate to any screen of their choice except in a linear format. But with the advent of web and the hypermedia structure, programs are now built with the feature such that the users can trace the path they like to within these computer based programs.

Learner control generally increases effectiveness, efficiency, and motivates learners. There are no disadvantages against using learner control as long as the control option does not confuse the learners. There are only arguments for and against the degree of learner control. Depover & Quintin (1992) mention that the degree of learner control depends on variables like age, previous knowledge, learning progress, complexity of material and familiarity with the subject. Some of the other variables that influence the degree of learner control are prior knowledge, student strategy and ability, learning progress, complexity of material and familiarity with the subject (Depover & Quintin, 1992; Hannafin, 1984; Milheim & Martin, 1991; Steinberg, 1989).

In computer based instruction, some of the control options learners have are, control over choice and sequencing of instruction, control over strategy and control over content. There has been research on learner control relating to choice of strategy such as level of difficulty of examples and practice items (Merrill, 1980; Kopcha, 2005). Chung & Reigeluth (1992) break down learner control into control over content, sequence, speed of learning, display or strategy, the internal process, and the advisory strategy. In this current study, the students had control over the sequence of instruction where the learners could choose the instructional elements in any sequence of their choice in the web based hypertext environment. Navigation option was provided as a menu bar at the top of the screen.

### **Purpose of Current Study**

In the previous studies conducted by the author (Martin, Klein & Sullivan, 2007; Martin & Klein, 2008), the learners did not have control on the sequence of the instruction in the program and had only linear navigation (program control). Hence in order to answer the questions of the effects of the instructional element practice with feedback when learners had control (non-linear navigation) and learners had no control (linear navigation), this study was



proposed. The IPSO instructional program with the same instructional content but with the changes to the navigation links was used in this study.

The purpose of this study was to investigate if a) the presence or absence of practice in a web based lesson had significant effect on student achievement, attitude and time b) if the navigation types which provided control over sequence of the instructional element (Linear, Non-Linear) had a significant effect on student achievement, attitude and time and c) if there was any interaction between practice and navigation type. The elements investigated in the study, practice with feedback, and linear and non-linear navigation type were combined into four different versions of web based programs in a manner that permitted investigation of the effectiveness of the program when practice was present and absent for both linear and non-linear navigation types.

The primary research questions for this study are listed below.

1. What is the effect of practice with feedback on achievement, attitude, and time when students use a web-based learning environment?
2. What is the effect of navigation type (linear and non-linear navigation) on achievement, attitude, and time when students use a web-based learning environment?
3. Does practice and navigation type (linear and non-linear navigation) interact to influence achievement, attitude and time?

The researchers anticipated that the combination of practice and linear navigation would have higher student achievement while practice with non-linear navigation would have higher student attitude.

## **Method**

### **Participants**

Participants were 240 freshman and sophomore undergraduate students enrolled in a computer literacy course at a large Southwestern University. 60 students participated in each treatment. The students enrolled in this course had varied background knowledge on computers and were from different majors including education, communication, journalism and others. Students participated in this study as part of the course requirement and the score in the post test was part of their course grade.

To avoid the variation in treatments with in the class (practice, no practice, linear navigation, and non-linear navigation), the students from were assigned to the treatments by classes and not by individual. The classes were randomly assigned to one of the four treatments based on the pretest average scores. It was a quasi-experimental study due to this nature of assignment to the treatments. This was one of the limitations to the study but helped to avoid differences in content, attitude or time spent on the program between the students enrolled in the same class.

### **Materials**

Four different versions of a web-based lesson on the topic *Input, Processing, Storage and Output of a Computer (IPSO)* were developed using Dreamweaver. IPSO explains the primary operations of the computer which are Input, Processing, Storage and Output. An introduction section was included before the primary operations were explained in detail. This section introduced what a computer is and classified it based on size, power and generation. It also explained the IPSO cycle. The next four sections described the concepts of the Input, Processing, Storage and Output operations in a computer and explained the function of the different components associated with that operation. The content used in this study was part of the required content for the course. The web-based lesson was pilot tested with five students before it was used in the study.

The four different versions of instructional program were as follows:

1. Program with Practice and using Linear Navigation (Program control)
2. Program with Practice and using Non-Linear Navigation (Learner Control over Sequence)
3. Program without Practice and using Linear Navigation (Program Control)
4. Program without Practice and using Non-Linear Navigation (Learner Control over Sequence)

In Linear navigation, the users go through the module in a linear path, page after page and this process can also be referred to as program control. Figure1 shows the linear or structure navigation approach that was used in the instructional material.

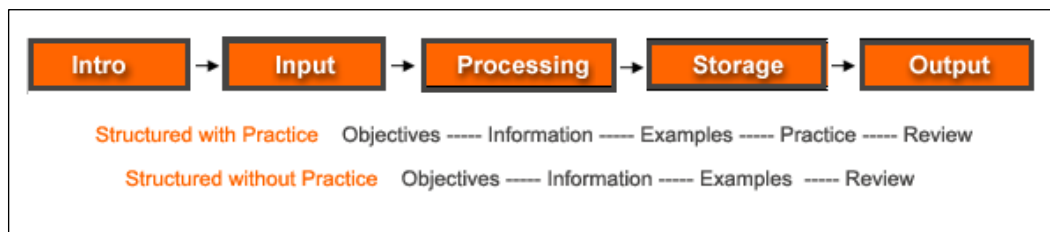


Figure 1. Linear Navigation or Program Control

And in Non-Linear navigation, users can navigate from one page to another in any particular order, both forward and backward and to any screen within the instructional program. This navigation type is also called as learner control over sequence. The instructional program in this study was built with Dreamweaver and the navigation type was hypertext links. Figure 2 shows the non-linear navigation or the unstructured navigation approach that used in the instructional material.

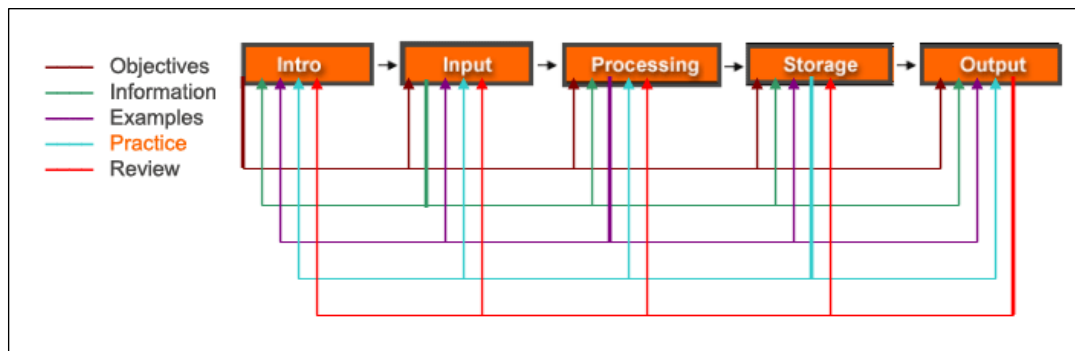


Figure 2. Non-Linear Navigation or Learner Control

## Know Your Computer Input

Intro	Input	Processing	Storage	Output	Posttest
INPUT	Objectives	Intro to Input	Input Devices 1	Input Devices 2	Practice Review

**Input Devices - Practice**

These practice questions are similar to the questions on the IPSO posttest. If you get a wrong answer while practicing, keep trying until you get it right.

1. A peripheral used to transfer data from the outside world into a computer system is called

- ☐ Input device
- ☐ Processing device
- ☐ Storage device
- ☐ Output device

Figure 3. Practice screen in the non-linear navigation program

### Practice Screens

The first two programs had practice screens in the program and it provided students with an opportunity to practice the content they were learning. There were a total of five practice screens, each of which contained five four-choice

multiple-choice questions. The student received immediate feedback after each response to a practice item. Students had the option to practice until they got the right answer. One practice screen was presented after each information screen. One example from the five items on a practice screen is shown in Figure 1 and Figure 2.

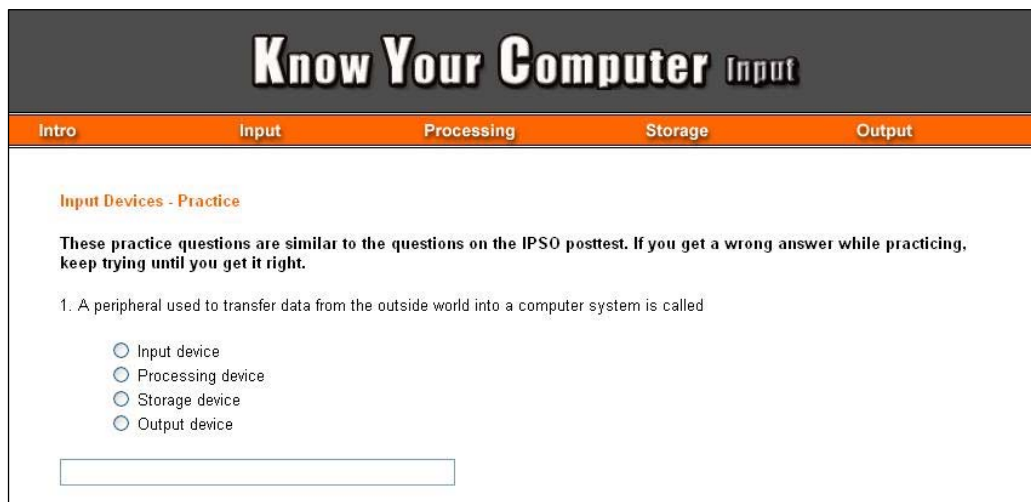


Figure 4. Practice screen in the linear navigation program

## Procedures

Eighteen sections of students ( $n = 240$ ) enrolled in the computer literacy course were blocked by classes and randomly assigned to the four treatment groups. The pretest, which took approximately 15 minutes to complete, was administered a week prior to the study. The classes were blocked into four groups based on their mean pretest scores, and one class within each block was assigned to each of the four treatments.

The participants used the web-based IPSO lesson during the sixth week of the semester. Participants met in a regular computer lab for instruction and were directed by the instructor to the web address for the instructional program. Each class was routed directly to its treatment version of the program. Students worked through the program at their own pace, averaging approximately one hour. Then they took the posttest and the attitude survey online. All six treatment groups followed the same procedure. Thus, the experimental differences in treatments occurred exclusively in the materials themselves and not in the procedure.

## Criterion Measures

The criterion measures consisted of a posttest and a student attitude survey. In addition, a pretest was used to assess subjects' knowledge of the content prior to the instruction.

*Pretest.* The pretest consisted of 20 multiple-choice questions covering the content with four response choice questions. A sample question that appeared on both the pretest and posttest is shown below.

*What is the purpose of a program counter?*

- a. Executes instruction
- b. Fetches instruction from the main memory
- c. Keeps track of successive instruction
- d. Temporary memory location

The overall mean score on the pretest was 8.38 or 42%, indicating that participants were not very knowledgeable about the content prior to instruction.

*Posttest.* The posttest consisted of the same 20 multiple-choice questions that were on the pretest. The reliability of the posttest was .65. The item analysis done on the posttest revealed that question 17 was the most difficult with a difficulty level of .42, question 3 was at .60, followed by question 1 and 2 at .62 and .63. The rest of the questions difficulty level varied between .70 and .99.

*Attitude Survey.* The attitude survey assessed student attitudes towards the instructional program and the presence or absence of the instructional events. The survey included 12 Likert-type questions that were rated strongly agree (scored as 4) to strongly disagree (scored as 0). The survey also included two open-ended questions that asked the participants what they liked best and least about the program. The survey was administered after the lesson and the posttest were completed. The reliability of the attitude survey was .83.

## Data Analysis

A 2X2 analysis of variance (ANOVA) test was conducted on data obtained from the achievement posttest and on the total time spent on the program. The attitude survey results for the Likert type items (Items 1-6) was analyzed using a 2X2 ANOVA. All analyses revealed significant differences.

## Results

### Achievement

The first research question investigated the effects of practice and navigation type on student achievement. Table 1 shows the mean scores and standard deviations for achievement on the pretest and posttest by treatment. The average pretest score was 8.46 (SD = 2.26) and posttest score was 15.91 (SD = 2.92). Participants who received practice and linear navigation scored the highest on the posttest (M=17.14) and those who received no practice and had non-linear navigation scored the lowest (M=14.78) on the posttest.

A 2X2 ANOVA conducted on the pretest data revealed no significant difference for practice main effect, navigation main effect or interaction. 2X2 ANOVA conducted between the treatment groups on the posttest revealed a significant practice main effect,  $F(1, 196) = 22.388, p < 0.01$ . Thus, there was a significant difference between the groups that received practice and no practice. Those who had received practice (M=16.84) performed significantly higher on the posttest compared to those who did not receive practice (M=14.98). However, there was no significant difference for those who had control over the instruction using the Linear navigation (M= 15.66) and Non-Linear navigation (M=16.16). There was also no significant interaction between practice and navigation type.

*Table 1. Means and Standard Deviations (SD) for Posttest scores by treatment*

Treatment	Pretest			Posttest		
	Linear Navigation Mean (SD)	Non-Linear Navigation Mean (SD)	Total	Linear Navigation Mean (SD)	Non-Linear Navigation Mean (SD)	Total
Program with Practice	8.24 (2.75)	8.50 (2.32)	8.37 (2.53)	17.14 (1.92)	16.54 (3.42)	16.84 (2.77)
Program without Practice	8.62 (2.60)	8.48 (2.21)	8.55 (2.40)	15.18 (2.81)	14.78 (2.77)	14.98 (2.78)
Total	8.43 (2.66)	8.49 (2.25)	8.46 (2.26)	16.16 (2.59)	15.66 (3.22)	15.91 (2.92)

### Attitude

The next research question dealt with the effects of practice and navigation type on student attitudes. Table 2 shows means for responses to the 6 Likert-type items on the attitude survey. The items were rated on a 5 point Likert scale from strongly agree (N= 4) to strongly disagree (N=0).

Table 2. Means for Attitude survey by treatment

Treatment	Program with Practice			Program without Practice		
Attitudes Items	Linear Navigation Mean (SD)	Non-Linear Navigation Mean (SD)	Total	Linear Navigation Mean (SD)	Non-Linear Navigation Mean (SD)	Total Mean (SD)
1. I learned a lot from this program.	3.32 (.587)	3.22 (.648)	3.27 (.617)	3.06 (.620)	3.12 (.659)	3.09 (.637)
2. The overall quality of the program was good	3.36 (.525)	3.26 (.600)	3.31 (.563)	3.26 (.600)	3.22 (.507)	3.24 (.553)
3. I would recommend this program to other students	3.22 (.648)	3.10 (.678)	3.16 (.662)	3.04 (.676)	2.94 (.586)	2.99 (.631)
4. I would enjoy using other computer programs like this one in future lessons	3.00 (.736)	3.04 (.755)	3.02 (.742)	2.96 (.735)	2.94 (.712)	2.95 (.720)
5. The program gave me enough opportunity to practice what I was learning	3.20 (.639)	3.26 (.694)	3.23 (.664)	2.56 (.760)	2.40 (.782)	2.48 (.772)
6. The program gave me enough control to move around the program.	3.22 (.616)	3.36 (.663)	3.29 (.640)	3.12 (.718)	3.30 (.544)	3.21 (.640)
Total	3.22	3.21	3.21	2.99	2.99	2.99

A MANOVA conducted on the overall attitude data revealed a significant difference on the 6 attitude questions,  $F(18, 535.06) = 4.33, p < 0.01$ . Follow-up 2x2 ANOVA conducted on the attitude data indicated significant differences for items 1(I learned a lot from this program) and 5 (The program gave me enough opportunity to practice) between the treatments who received practice and those who did not. No items showed significant difference for the navigation type (linear versus non-linear navigation). There was no interaction between practice and navigation type on the attitude data. For both the above items that had practice main effect, the participants who received practice had rated it significantly higher than those who did not receive practice.

The attitude data also showed that participants who used the program with practice had higher attitudes and had rated higher on all the six items compared to those who did not receive practice. Item 2 “The overall quality of the program was good” was the rated the highest ( $M = 3.28$ ) by both the practice and no practice group. Those who did not receive practice recognized the absence of practice in their programs and rated item no 5 “The program gave me enough opportunity to practice what I was learning”, the lowest ( $M = 2.48$ ). There weren’t much difference in attitudes of the participants when comparing the navigation method they received and were almost equivalent in their ratings based on navigation.

The attitude survey also included two open-ended questions that asked the participants what they liked the best and least about the program. The four most frequent responses for what participants liked the best in the program were (1) the practice questions ( $n=65$ ) (2) clear navigation and structure ( $n=45$ ) (3) the review section ( $n=43$ ) (4) graphics, animations and visuals ( $n=34$ ) (4) Highly informative ( $n=30$ ). The most frequent response for what parts liked the least were (1) Very long program ( $n= 36$ ) (2) Lot of information ( $n=23$ ).

## Time

On calculating the time spent in the program (table 3), there were no significant differences between the groups based on practice and navigation. Those who received practice spent more time on the program ( $M = 35.56$ ) than those who did not receive practice ( $M = 31.33$ ), but there were no significant differences. Both the navigation types, linear ( $M=33.49$ ) and non-linear ( $M=33.40$ ) spent about the same amount of time in the program.

Table 3. Means and Standard Deviations (SD) for Time spent in minutes by treatment

Treatment	Time		
	Linear Navigation Mean (SD)	Non-Linear Navigation Mean (SD)	Total
Program with Practice	35.98 (13.59)	35.14 (18.84)	35.56 (16.34)
Program without Practice	31.00 (13.15)	31.66 (16.19)	31.33 (14.68)
Total	33.49 (13.53)	33.40 (17.56)	33.44 (15.64)

## Discussion

This study examined the effects of Practice and Navigation type (Linear and Non-Linear) on achievement, attitude and time. College students enrolled in a computer literacy course used a web-based lesson delivered on the web to learn about input, processing, storage and output of a computer (IPSO). The computer based lesson included multiple choice practice items and two types of navigation (Linear, Non-Linear). Linear navigation directed them from one page to the next where as in Non-Linear navigation they had the freedom to navigate in any path. They had control over the sequence of instruction.

Results indicated that there was a significant difference between the groups that received practice and no practice, but there was no significant difference in Linear and Non-Linear navigation. There was no significant interaction between practice and navigation type.

## Achievement

Practice resulted with a significant difference both in achievement and in attitudes. Practice provides an opportunity for learners to confirm their correct understanding, and repetition increases the likelihood of retention of new knowledge (Kruse & Kevin, 1999). In this web-based lesson, participants who received practice were also given feedback which confirmed the student's answer as being correct or indicated that it was incorrect. This feedback strengthens the probability of correct responses and reduces the probability of subsequent incorrect responses. (Philips et al., 1988). The confirmation of the answer during practice increased the likelihood of retention of the content. Presence of practice results in interaction between the lesson and the learner. It is effective in performance when it is aligned with the assessment in the form of a posttest and with the skills, knowledge and attitudes reflected in the objectives (Reiser & Dick, 1996). In this web-based lesson, practice was directly aligned with the posttest and with the information presented. The findings of the current study is consistent with previous research on computer based instruction that found practice had an effect on learning (Hannafin, 1987; Philips et al., 1988).

Navigation types (linear, non-linear) did not result in a significant difference. This could have been due to the fact that even though the non-linear treatment had the flexibility to take any path that they decided, the computer based lesson was well structured and organized. The lesson was designed using all the instructional elements and was instructionally aligned from objectives to assessment. The results could have been different, if the instructional material was not well structured and did not have the other instructional elements or instructional alignment. Though there was no significant difference, participation of the linear navigation program scored higher than the participants of the non-linear program.

Participants who received practice and linear navigation scored the highest on the posttest and those who received no practice and had non-linear navigation scored the lowest on the posttest. When enough of structure is provided and the instructional material is well designed, the students do better when the navigation is linear and they are forced through every screen where they learn from every instructional element such as objective, practice, feedback and review. The absence of the instructional element practice and non-linear navigation resulted in the lowest posttest scores.

Presence of practice not only diminishes the importance of other instructional elements such as objectives (Hannafin, 1987; Hannafin et al, 1987; Philips et al, 1988) but also diminishes the importance of other non-instructional elements such as navigation. Practice was an instructional element which was directly aligned to the objectives and posttest and had an effect on student achievement, and navigation was the non-instructional element and did not have an effect on student achievement.

## **Attitude**

The attitude survey had 6 Likert-type items. Participants who received practice in their computer lessons had higher attitudes compared to those who did not receive practice. Those who received practice had an opportunity to interact with the web-based lesson and it helped them perform better and have higher attitudes. Higher attitudes of the participants who received practice could have been due to the feedback they received during practice, which strengthened the probability of correct responses and reduced the probability of subsequent incorrect responses. The attitude results remained consistent with the achievement results.

Item 1 “I learned a lot from this program” and Item 5 “The program gave me enough opportunity to practice what I was learning” resulted with significant differences between the treatments which received practice and which did not receive practice. The significant difference on item 5 (The program gave me enough opportunity to practice what I was learning) reveals that the participants realized the presence of practice in their computer based lessons. And also on item 1 (I learned a lot from this program ), those who received practice had higher attitudes that they had learned a lot from the program. The presence of interactive well aligned practice items, which provided feedback and corrected their understanding of the concepts, must have been the reason for them to state that they had learned a lot from this program. Item 6 “The program gave me enough control to move around the program” which is about the navigation aspect of the program did not result in significant differences. Thus both in attitudes and achievement, there was significant differences for the presence of practice but not for navigation type.

In the open ended section, practice topped the list for what the participants liked the best about the program and was followed by the clear structure and navigation. It can be noted that participants realized that practice made a difference in the program. We may not have had a significant difference in the navigation types, but from the open ended question it is shown that students were aware of the structure and navigation used in the program and had rated it as the second best feature in the program.

## **Time**

Time did not result in any significant difference for both practice and navigation type. Though the participants in the practice treatment spent longer time in the program, it was not significantly different from the treatment who did not receive practice. But the time spent by both the navigation types were almost the same. Participants in the linear navigation treatment spent (M=33.49) minutes and those in the non-linear navigation treatment spent (M=33.40) minutes. This shows that the different navigation types did not matter in regard to time spent.

## **Conclusion**

This study has once again reinforced the importance of practice in web based/computer based lessons. It once again confirmed Practice to be effective in help in student achievement and attitude. If the lesson is instructionally sound with important instructional elements such as practice then irrespective of the non-instructional element such as navigation type, learners have high achievement and attitudes. It also showed that when the lesson is well structured then the effect of navigation is not seen.

This study has implications to the design and development of web-based instructional material. Practice is an effective instructional element for enhancing student achievement. This suggests that it should be included in web-based instruction especially when students are tested using items aligned with the objectives and practice items. We also recommend including different types of practice items. In this study, multiple choice practice items with immediate feedback to students were included. Also, the content was more on learning facts and concepts. Future

research should focus on variation in the instructional content and the type of practice and feedback involved. Additional research should examine how instructional elements in computer-based instruction influence outcomes such as problem solving and complex learning tasks. As was done in this study, research in these settings should include measures of student achievement, attitudes and time. Further studies can be conducted to test the effectiveness of the other instructional elements such as objectives, examples, and review along with the different non-instructional elements such as intelligent agents and other usability elements including navigation types. It will be helpful to measure the practice scores and their correlation to the posttest scores. Studies of this nature will continue to inform designers about the influence of instructional elements on learning and performance.

## References

- Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the Impact of an Animated Agent's Voice. *Contemporary Educational Psychology*, 30, 117-139
- Boechler, P.M., & Dawson, M.R.W. (2002) Effects of navigation tool information on hypertext navigation behavior: a configural analysis of page-transition data. *Journal of Educational Multimedia and Hypermedia*, 11, 2 2002. 95-115
- Chung, J. & Reigeluth, C.M. (1992). Instructional prescriptions for learner control. *Educational Technology*, 32, 10, 14-20
- Clark, R. C., & Mayer, R. E. (2007). Using rich media wisely. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and Issues in Instructional Design and Technology* (2<sup>nd</sup> Ed.). Upper Saddle River, NJ: Pearson/Merrill-Prentice Hall.
- Dee-Lucas, D., & Larkin, J.H. (1995). Learning from electronic texts: Effects of interactive overviews for information access. *Cognition and Instruction*, 13, 3, 431-468.
- Delgado, A. R. & Prieto, G. (2003). The effect of item feedback on multiple-choice test responses. *British Journal of Psychology*, 94, 73-85.
- Depover, C. E. & Quintin, J. J. (1992). Learner control versus computer control in professional training context. In M. Giardina (ed.), *Interactive Multimedia Learning Environments: human factors and technical considerations on design issues*. Berlin, Springer-Verlag, 234-247.
- Dias, P., & Souza, A.P. (1997). Understanding navigation and disorientation in hypermedia learning environments. *Journal of Educational Multimedia and Hypermedia*, 6, 2, 173-185.
- Dick, W., Carey, L. M., & Carey, J. O. (2005). *The Systematic Design of Instruction*, (6th ed.). Boston: Pearson/Allyn and Bacon.
- Freitag, E & Sullivan, H. (1995). Matching Learner Preference to Amount of Instruction: An Alternative Form of Learner Control. *Educational Technology Research and Development*, 43, 2, 5-14
- Forcier, R. C., & Descy, D. E. (2002). *The computer as an educational tool*, (3<sup>rd</sup> Ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Foshay, W. R., Silber, K. H., & Stelnicki, M. B. (2003). *Writing training materials that work*. San Francisco: Jossey-Bass/Pfeiffer.
- Gagné, R. (1965). *The Conditions of Learning* (1<sup>st</sup> ed.). New York: Holt, Rinehart & Winston.
- Gagne, R. (1985). *The Conditions of Learning* (4th ed.). New York: Holt, Rinehart & Winston .
- Gagne, R.M. (1988). Mastery Learning and Instructional Design. *Performance Improvement Quarterly*, 1, 1
- Gagne, R., Briggs, L. & Wager, W. (1992). *Principles of Instructional Design* (4th Ed.). Fort Worth, TX: HBJ College Publishers.
- Gagne, R. M., Wager, W.W., Golas, K.C., & Keller, J.M. (2005). *Principles of Instructional Design*. (4<sup>th</sup> Ed.) Belmont, CA: Wadsworth/Thompson Learning.
- Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development* 7(3), 6-10



- Hannafin, M. (1987). The effects of orienting activities, cueing and practice on learning of computer based instruction. *Journal of Educational Research*, 81(1), 48-53
- Hannafin, M., Philips, T., Rieber, T., & Garhart, C. (1987). The effects of orienting activities and cognitive processing time on factual and inferential learning. *Educational Communications and Technology Journal*, 35, 75-84
- Hammond, N & Allinson, L. (1989) Extending Hypertext for Learning: An Investigation of Access and Guidance Tools. A. Sutcliffe and L. Macaulay(Eds.) People and Computers V, Cambridge: Cambridge University Press, 293-304
- Klein, J.D & Pridemore, D.R. (1994) Effects of Orienting Activities and Practice on Achievement, Continuing Motivation, and Student Behaviours in a Cooperative Learning Environment. *Educational Technology Research and Development*, 42, 4, 41-54
- Klein, J. D., Spector, J. M., Grabowski, B., & de la Teja, I. (2004). *Instructor competencies: Standards for face-to-face, online, and blended settings*. Greenwich, CT: Information Age Publishing.
- Kruse, K. & Kevin, J. (1999).Technology-based Training: The Art and Science of Design, Development and Delivery. Jossey Bass publication
- Kopcha, T.J. (2005). Learner preferences, prior knowledge, and learner control over practice and examples in computer-based instruction. Paper presented at the Association of Educational Communications and Technology.
- Ku, H.Y & Sullivan, H.J (2002) Student performance and attitudes using personalized mathematics instruction. *Educational Technology Research and Development*, 50, 1, 21-34
- Kulhavy, R.W. (1977) Feedback in written instruction. *Review of Educational Research*, 47, 1, 211 - 232
- Kulhavy, R., White, M., Topp, B., Chan, A., & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, 10, 285- 291
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1, 279-308.
- Mager, R. F. (1997). *Preparing Instructional Objectives* (3<sup>rd</sup> Edition). Center for Effective Performance: Atlanta, GA.
- Martin, F., Klein, J.D. & Howard, S.. (2007). The impact of instructional elements in computer-based instruction. *British Journal of Educational Technology*, 38, 4, 623 - 636
- Martin, F., & Klein, J.D. (2008). Effects of Objectives, Practice, and Review in Multimedia Instruction. *Journal of Educational Multimedia and Hypermedia*, 17 (2), 171-189
- Mattiske, C. (2001). *Train for results: Maximize the impact of training through review*. Warriewood, Australia: Business and Professional Publishing.
- McDonald, S., & Stevenson, R. (1998). Spatial versus conceptual maps as learning tools in hypertext. *Journal of Educational Hypermedia and Multimedia*, 8,1, 43-64.
- Merrill, M.D. (1980). Learner control in computer based learning. *Computers and Education*, 4, 77-95.
- Merrill, M. D. (2002) First principles of instruction. *Educational Technology Research and Development*, 50(3), 43-59.
- Milheim, W. D., & Martin, B.L. (1991). Theoretical bases for the use of learner control: three different perspectives. *Journal of Computer-Based Instruction* 18(3), 99-105.
- Morrison, G.R., Ross, S.M., & Kemp, J.E. (2006). *Designing effective instruction* (5<sup>th</sup> Ed.), New York: John Wiley.
- Philips, T., Hannafin, M., & Tripp, S. (1988). The effects of practice and orienting activities on learning from interactive video. *Educational Communication and Technology*, 36, 93-102.
- Reiser, R. A. (2002). A history of instructional design and technology. In Reiser, R. A., & Dempsey, P. (Eds.), *Trends and Issues in Instructional Design and Technology*, Upper Saddle River, NJ: Merrill-Prentice Hall, 26-53.
- Reiser, R. A. & Dick, W. (1996). *Instructional Planning: A guide for teachers* (2<sup>nd</sup> Ed.), Boston: Allyn and Bacon.

- Richey, R. (2000). *The Legacy of Robert M. Gagné*, Syracuse, NY: Eric Clearinghouse on Information and Technology.
- Schankenberg, H.L., Sullivan, H.J., Leader, L., & Jones, E.K. (1995) Matching learner preference to preferred amount of Instruction. *Paper presented at the Annual Convention of the Association for Educational Communications and Technology*, Indianapolis, IN, USA.
- Shneiderman, B. & Kearsley, G. (1989) *Hypertext Hands On! An Introduction to a New Way of Organizing and Accessing Information*, Reading, MA: Addison Wesley.
- Steinberg, E. R. (1989). Cognition and learner control: a literature review, 1977-88. *Journal of Computer-Based Instruction*, 16 (4), 117-24.
- Su, Y., & Klein, J. (2006). Effects of Navigation Tools and Computer Confidence on Performance and Attitudes in a Hypermedia Learning Environment. *Journal of Educational Multimedia and Hypermedia*. 15 (1), 87-106.
- Smith, P.L., & Ragan, T.J. (2005). *Instructional design* (3<sup>rd</sup> Ed.), Hoboken, NJ: John Wiley.
- Sullivan, H. J. & Higgins, N. (1983). *Teaching for competence*, New York: Teachers College Press.

## Content Recommendation Based on Education-Contextualized Browsing Events for Web-based Personalized Learning

Feng-Hsu Wang

Department of Computer Science and Information Engineering, Ming Chuan University, Taiwan // fhwang@mail.mcu.edu.tw

### ABSTRACT

The WWW is now in widespread use for delivering on-line learning content in many large-scale education settings. Given such widespread usage, it is feasible to accumulate data concerning the most useful learning experiences of past students and share them with future students. Browsing events that depict how past students utilized the learning content to accomplish higher levels of achievement are especially valuable. This paper presents a new method for identifying potentially effective browsing events based on a contextualized browsing model built through association mining and statistical techniques. The model annotates browsing events with several contextual factors, including educational ones (group relevance and performance relevance) and non-educational ones (support and confidence). Based on this model, a personalized content recommender was implemented in a Web-based learning content management system, called IDEAL, to deliver personalized learning content based on a student's browsing history. An experiment was conducted to compare the user feedback concerning the recommendations provided through different recommendation models. The results show that students with different levels of achievement prefer different types of contextualization information. Finally, another performance experiment demonstrated that the contextualized browsing model is more effective in improving learning performance than the pure association mining model.

### Keywords

Web-based learning, Personalized learning, Content annotation, Association mining, Content recommendation

### Introduction

The WWW contains a huge amount of learning content and is now in widespread use for delivering on-line learning content in many large-scale education settings. However, the huge amount of learning content has presented a problem to on-line students. Students are apt to get lost in the huge content space. Therefore, personalized learning becomes an important mechanism of a learning system that can guide students by automatically recommending learning content to their needs in a just-in-time manner (Zaiane, 2001).

The target of personalized learning varies with the types of learning needs it is designed to (McNaught, Kennedy & Majoor, 2002). For example, for just-in-time training, the main focus is on delivering appropriate information that workers need to solve problems, perform specific tasks or update their knowledge and skills. To perform such a type of personalized learning, specific knowledge about the task structure is required, such as in AIMS (Aroyo & Dicheva, 2001). On the other hand, for Web-based learning where students build the domain knowledge by studying learning content and navigating through a rich set of learning resources, personalized recommendation focuses on providing the next-step browsing suggestions so that students can build knowledge effectively with no disorientation in the learning environment (Kinshuk & Lin, 2003). This type of personalized content recommendation is the main focus of this research.

As to the design of personalized content delivery platforms, a new eLearning application, called Learning Content Management System (LCMS) (Brennan, Funke & Anderson, 2001), has been developed as a critical component of the personalized learning paradigm in which the emphasis is shifted from the knowledge of the instructor to the knowledge inherent in the content. Besides, a significant change is taking place in the way on-line learning is going in recent years. This change is the advocacy of learning objects to support personal learning needs. The concept of learning objects features the reusability of learning resources in whatever contexts they can be applied (Mohan, 2004; Mphan & Greer, 2003). Based on the key idea of LCMS that it will deliver what is needed at the time when it is needed, personalized content recommendation has become a very important learning support in LCMS (Brusilovsky & Vassileva, 2003; Denaux, Dimitrova & Aroyo, 2004).

One approach to personalized content recommendation is knowledge-oriented, which is based on three types of knowledge: domain ontology, content knowledge and student models (Papanikolaou & Grigoriadou, 2003; Sampson, Karagiannidis & Cardinali, 2002; Mittal, Krishnan & Altman, 2006). Domain ontology plays a shared language base for representing content knowledge and student models. A knowledge-based recommender could give highly individualized content recommendations. However, this approach incurs the cost of developing knowledge bases and is limited by the extent to which the student model is accurate.

Another approach to personalized content recommendation is to acquire knowledge about effective learning experiences and then share the knowledge with future users (Mulvenna, Anand & Buchner, 2000). As far as content recommendation is concerned, the ability of a web-based learning system to keep traces of students' browsing behavior can be exploited to promote the learning paradigm of learning by other's learning experiences (Najjar, Wolper & Duval, 2006). For those online courses that operate in a yearly cycle, which are common in most universities, the browsing traces can be accumulated and enriched year by year. With an appropriate analysis of these traces, navigation patterns can be discovered and shared with future students (Liang & Leifer, 2000; Najjar, Wolper & Duval, 2006).

There have been several research efforts to clarify the relationship among student cognitive characteristics, navigation patterns and learning performance. Several research results show that student cognitive characteristics seem to have a great effect on the strategies of how they learn, navigate and search learning resources (Liu & Reed, 1995; Ford & Chen, 2000; Chen & Macredie, 2002). However, there is little evidence showing that the differences in cognitive characteristics have a great impact on learning performance. Some research results show there is a positive relationship between cognitive styles and learning performance in hypermedia environments (Andris, 1996; Parkinson & Redmond, 2002); while others failed to find a connection between them (Liu & Reed, 1994; Calcaterra, Antonietti & Underwood, 2005). In particular, Calcaterra et al (2005) showed that stylistic differences were not associated with navigation patterns, and specific browsing behavior induced a better learning outcome. In short, they found that specific browsing behavior did have an effect on learning performance, but such behavior is not influenced by student cognitive styles. While there is evidence that individual factors do affect hypermedia navigation, their impact on learning performance is not simple. Based on the findings of Calcaterra et al (2005), it is assumed in this paper that sharing good browsing behavior is beneficial to students regardless of their cognitive characteristics.

Nevertheless, the tremendous amount of browsing traces arouses the need of an automatic pattern extraction method. As to acquiring access patterns from historical traces, association mining (Agrawal, Imielinski & Swami, 1993; Agrawal & Srikant, 1994; Chen, Han & Yu, 1996) has been applied successfully in commercial applications. Furthermore, there is also research in reusing browsing patterns in the educational context (Mobasher, Cooley & Srivastava, 2000; Wang & Shao, 2004), where association mining is adopted as well.

The rationale behind association mining is that frequent association events can be used to anticipate the potential needs of users. However, several problems arise when a pure association mining approach is applied in the educational context. For example, in many researches, only the frequency and confidence of association events are considered to support a user's decisions on taking which recommendations. Besides, the 0/1 transactional model of association mining does not distinguish between browsing events with different amounts of efforts that past students have made in pursuing their learning goals; nor does it care about who have generated these events. It is our belief that past browsing patterns need to be contextualized appropriately in an educational setting before they can be used effectively by future students.

This paper presents a new approach to reusing past browsing patterns based on a contextualized browsing model. Specifically, a model is proposed to contextualize browsing events with two more educational factors: group relevance and achievement relevance, in addition to the support and confidence factors adopted in pure association mining. The group-relevance factor is concerned with the relation between browsing events and student groups of different achievements, while the achievement-relevance factor has to do with the correlation between browsing events and learning performance. By contextualizing the browsing events based on this model, browsing patterns can be shared with the students by providing them with the different educational perspectives which they can use to make a decision on what to study next.

Based on the contextualized browsing model, a content recommender was implemented in a Web-based learning content management system (LCMS), called IDEAL (Integration and Dissemination of Electronic Assessment and Learning). Finally, an experiment was conducted to compare the user feedback concerning the recommendations provided through different recommendation models. The results show that students with different levels of achievement prefer different types of contextualization information. Furthermore, another performance experiment demonstrated that the contextualized browsing model is more effective in improving learning performance than the pure association mining model.

## Organization of the paper

The rest of this paper is organized as follows. The second section depicts the connection between association mining and personalized learning. The third one introduces the contextualization process for browsing patterns. Next, a brief introduction of how the recommender interacts with the IDEAL system is given, followed by a description of the evaluation of the system as well as a discussion on the experimental results. At last, a summary and future research is given.

## Association Mining and Personalized Learning

Association mining is one of the most well studied methods in data mining (Agrawal, Imielinski & Swami, 1993; Agrawal & Srikant, 1994; Chen, Han & Yu, 1996). It has served as a useful tool for discovering correlated items in a large transaction database. It explores the probability of the event that when certain items are present in a transaction, what other items would also present themselves in the same one. In effect, an association rule is a statement of the form  $X \rightarrow Y$ , where  $X$  and  $Y$  are two disjoint item sets. An interpretation of such an association rule in a business context is that if a customer has bought the items in  $X$ , he/she would like to buy the items in  $Y$  in the same transaction

When it comes to measurement of effectiveness of association rules, support and confidence are the most adopted. The support value indicates the occurrence frequency of an item set in a transaction database. It is defined as the ratio of the occurrences of an item set over the total database size. For example, suppose there are  $x$  transactions containing items in an item set  $X$ , and the database has a total of  $y$  transactions, then the support of the item set  $X$  would be denoted by  $\text{sup}(X) = (x/y)$ . Furthermore, in the minimum support approach, a set of items  $X$  is called a large itemset if the support rate of  $X$  meets a minimum support requirement. Association rules are generated only by a large itemset that has a support value greater than or equal to the minimum support. For an association rule  $X \rightarrow Y$ , the support of the associated pattern  $X \cup Y$  is used to check whether the rule is significant or not, implying that it is a frequent transaction pattern. On the other hand, confidence of an association rule  $X \rightarrow Y$  represents the strength of the implication of the rule. It is defined as an estimation of the conditional probability computed by  $\text{sup}(X \cup Y) / \text{sup}(X)$ . An association rule is said to be more reliable if it has a higher confidence.

From the educational perspective, frequent itemsets can be browsing events consisting of learning objects often browsed together in a learning environment. The association rule corresponding to a browsing pattern can be used to generate personalized recommendations as follows. For example, let's consider a browsing pattern  $(LO_1, LO_2, LO_3, LO_4)$  with support 0.05, and an association rule  $(LO_1, LO_2, LO_3) \rightarrow (LO_4)$  with confidence 0.95, saying that 95% of the browsing transactions containing  $LO_1, LO_2$  and  $LO_3$  would also contain the learning objects  $LO_4$ . According to this rule, once a student has browsed learning objects  $LO_1, LO_2$  and  $LO_3$ , it would be anticipated that the student would like to browse  $LO_4$  as well (with confidence 0.95). Once the student adopt the recommendation, it can be said that s/he is following the past browsing pattern  $(LO_1, LO_2, LO_3, LO_4)$ .

Note that such a mechanism for personalized learning is different from those knowledge-based ones (Papanikolaou & Grigoriadou, 2003; Sampson, Karagiannidis & Cardinali, 2002). The essence of this type of personalized learning is the herd mentality exploited through emphasizing the popularity of learning objects that are often browsed together. It is based on the assumption that students often feel like they have made the right choices when they are doing as those students having similar need did.

Furthermore, there are differences between applying association mining to the educational arena and to the commercial one. In the educational context, association rules resulting in patterns with high support values may represent learning trends, while those with low support values but high confidence values may represent “special” ones. Unlike those commercial applications that focus only on events with large support values, it appears that both event categories, either trends or special ones, might expose valuable information to instructors and students. Besides, there are limitations of the 0/1 transactional model of pure association mining. It does not distinguish between transactions containing different amounts of items; nor does it care about who have generated those transactions. It is our belief that experiences from past students need to be contextualized appropriately before they can be used effectively by future students. Therefore, some contextualization task is required to help students better understand the rationale behind the system’s recommendation and make a more cohesive decision.

## The Contextualized Browsing Model

In this research, the browsing model that depicts past browsing behavior is developed based on frequent browsing events. In addition to the non-educational contextual information of *support* and *confidence* in pure association mining, browsing events are provided with more educational contextual information, that is, the group relevance and achievement relevance. Figure 1 shows the data analysis process that integrates association mining and statistical techniques to extract contextualized browsing patterns that are useful for personalized delivery of learning objects.

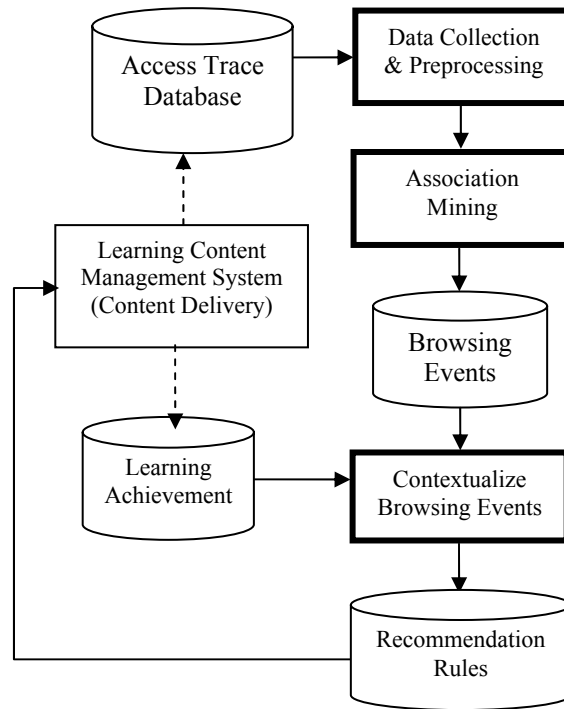


Figure 1: A data analysis process for contextualizing browsing events through data mining and statistic techniques

### Data collection and preprocessing

During the progression of an online course, the content access traces associated with each learning activity need to be collected so that we can identify what students have browsed to achieve the goal of the corresponding learning activity. The learning environment is equipped with content access tracking to facilitate the data collection task. A proper design of data tracking needs to select appropriate tracking attributes such as user ID, learning activity ID, content type and content ID, the date and session during which it was accessed, and so on. Table 1 shows the format of access traces that are saved in a database.

The collected data has to be cleaned up and transformed properly before it can be analyzed further. In situations where user sessions are not easy to identify, several heuristic methods have been explored to decide on user sessions from the logged access history (often logged in an http server) (Chen, Han & Yu, 1996). Nevertheless, in our situation where login actions are required from registered users, the problem of session determination is trivial. A number of preprocessing steps are performed to prepare the data for further analysis. For example, all browsing records are sorted in an ascendant order with the user ID as the major key and the event time as a minor key. After this sorting step, it is easy to identify user sessions by packing contiguous records that follow a *login* record until the next one. Specifically, browsing records picked out between two successive login records are grouped into a browsing session. Notice that a web trail may be followed just because it seems attractive for students, and that doesn't necessarily mean it is a right choice from a student's perspective. To reduce the trail-attraction effect, this research adopts the minimum page-stay session-grouping approach (Wang and Shao, 2004) that filters out learning objects browsed by a student in less than a specified time (say 7 seconds).

*Table 1: The format of access traces.*

Data Field	Description
<i>User ID</i>	Identifier of students.
<i>Activity ID</i>	Identifier of a learning activity
<i>LO ID</i>	Identifier of the referenced learning objects (assets, SCOs).
<i>Activity Type</i>	Activity type such as <i>login</i> , <i>browsing</i> , and so on.
<i>Stay Time</i>	Stay time of the LO access.

### Discovering browsing events through association mining

In this step, association mining is adopted to discover the set of browsing events consisting of learning objects *frequently* browsed together in a learning session. Recommendation rules can then be generated based on these browsing events. In this research, we adopt the one-page recommendation rule formation, in which a rule has multiple antecedents but only one consequent. The antecedents represent learning objects currently browsed, and the consequent represents the learning object for recommendation. Each recommendation rule is tagged with two labels representing a partitioned range of the support and confidence values, respectively. The label for support is assigned one of the five terms: *very-infrequent*, *infrequent*, *moderately-frequent*, *frequent*, *very-frequent*. Similarly, the confidence label is assigned one of the five terms: *very-uncertain*, *uncertain*, *moderately-certain*, *certain*, *very-certain*. The two labels are provided to help students understand the recommendations in terms of their frequency and certainty.

### Contextualization of browsing events

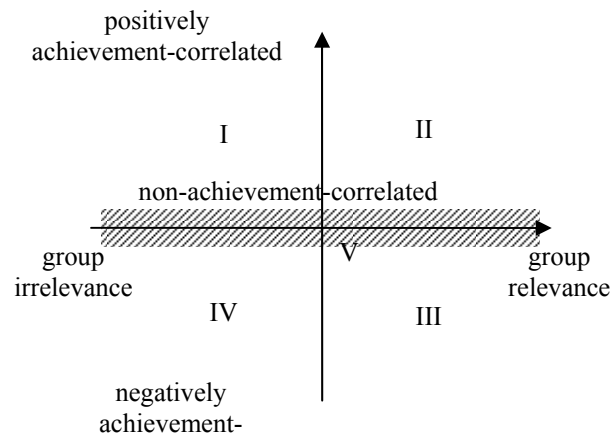
Additional contextualization steps are performed on browsing events using the educational contextual information of group relevance and achievement relevance. The former is concerned with how a browsing event relates to student groups of different learning achievements. For example, while some browsing events may receive more attention from students who are high achievers, they might not receive as much attention from students who are low achievers. On the other hand, there might be some browsing events that receive attention from all students, regardless of their learning achievements. Knowing this kind of knowledge is helpful for personalized learning because it could uncover browsing patterns specific to different student groups.

The second contextualization step is concerning how a browsing event correlates with learning achievement. Though the correlation is not necessarily a causality that guarantees effective learning, it reveals a potential direction for utilizing learning objects effectively. As to those browsing events showing a positive correlation with learning achievement, they reveal the potentially significant learning objects that students should pay more attention to. On the other hand, for those browsing events showing a negative correlation with learning achievement, they might indicate a potential problem with students trying to learn from the learning objects of the browsing events. Finally,

for those browsing events showing no significant correlation with learning achievement, it can only be assured that those events do not play significant roles in making learning achievement different.

### *The model*

In summary, the two-dimensional contextualization model can be depicted as shown in Figure 2. The group-relevance dimension is concerned with the browsing characteristics specific to student groups, while the achievement-relevance dimension concerns the correlation between browsing events and learning achievement.



*Figure 2: The browsing-event model with educational contextual information*

According to Figure 2, there are five types of browsing events. For type I events, they are not characteristic of any student group, but they have positive correlation with learning achievement. These events are significant because students who pay more attention to them are apt to perform better.

As to type II events, they are characteristic of some student group and have a positive correlation with learning achievement. These events are significant because a specific group of students paying more attention to them are apt to learn better. Therefore, the corresponding learning objects are good recommendation candidates for future students of that group.

For type III events, they are characteristic of some student group, but they have a negative correlation with learning achievement. These events are significant because students of that group are apt to perform worse when they pay more attention to the events. One possible reason for this might be that students of that specific group have problems learning from the corresponding learning objects. Therefore, more remedial instruction based on those learning objects should be given to the students in that group.

For type IV events, they are not characteristic of any student group, and they have a negative correlation with learning achievement. These events are significant because students who pay more attention to them are apt to perform worse. Therefore, more remedial instruction based on those learning objects should be given to all students in the class.

Finally, type V events, as marked in the shaded area of Figure 2, are not achievement-correlated. They are less significant in terms of the contextual factors considered in this paper. Nevertheless, this does not mean that the corresponding learning objects are not effective for learning. Instead, they still provide some information of past browsing behavior, and this type of information should also be referenced by future students for a general learning purpose.



## The method

To cope with the 0/1 transaction problem of pure association mining, the degree of *attention* paid by students to browsing events needs to be quantified in some way. This research introduces a concept of *session browsing effort* (SBE) that a student has made with respect to a browsing event  $e$  in a learning session  $s$ . Specifically, it is defined as a summation of occurrences of each learning object of  $e$  in session  $s$ . Note that the browsing effort of an event  $e$  in a session  $s$  is meaningful only if the session  $s$  contains the pattern  $e$ . For example, given a browsing event  $e = \{LO_1, LO_2, LO_3\}$ , and a session  $s_1 = \langle LO_1, LO_4, LO_1, LO_2, LO_3, LO_5 \rangle$ , the browsing effort with respect to event  $e$  in session  $s_1$  is  $2+1+1=4$ , because there are two visits of  $LO_1$ , one visit of  $LO_2$ , and one visit of  $LO_3$  in session  $s_1$ . On the other hand, in a session  $s_2 = \langle LO_1, LO_4, LO_1, LO_2, LO_5 \rangle$ , the browsing effort with respect to event  $e$  in session  $s_2$  would be 0 since session  $s_2$  does not contain  $LO_3$  in event  $e$ . Finally, the *total browsing effort* (TBE) of a student with respect to an event  $e$  is defined as the total summation of all session browsing efforts (SBEs) of the student with respect to the event  $e$ . The quantification of browsing events by a measure of object visits is in line with Yin (2001) and Calcaterra et al (2005), who showed that higher performance was associated with re-visiting learning objects.

Now, we define the “mean browsing effort” (MBE) of a student group with respect to a browsing event as the average TBE of all students in that group. Those events that have significant difference in MBE among different student groups are called “group-relevant” browsing events; otherwise they are called “group-irrelevant” ones.

To identify group-relevant browsing events, students are first divided into three groups according to their learning achievements (often measured by their post-test scores) after a learning activity. The first group consists of the top 25% students, which is called the “high-achievement” group. The second consists of the next 50% students, which is called the “middle-achievement” group. The last one is called the “low-achievement” group that covers the bottom 25% students. The relevance of a browsing event to student groups is then tested by conducting a one-way ANOVA  $F$  test on the mean browsing efforts (MBEs) among the three student groups. However, when the data does not meet the normality assumption of an ANOVA, the non-parametric Kruskal-Wallis Test (Siegel & Castellan, 1988) is performed instead. When a group-relevant event is identified, the t-test is performed on each pair of the student groups to identify the group relevant to the event.

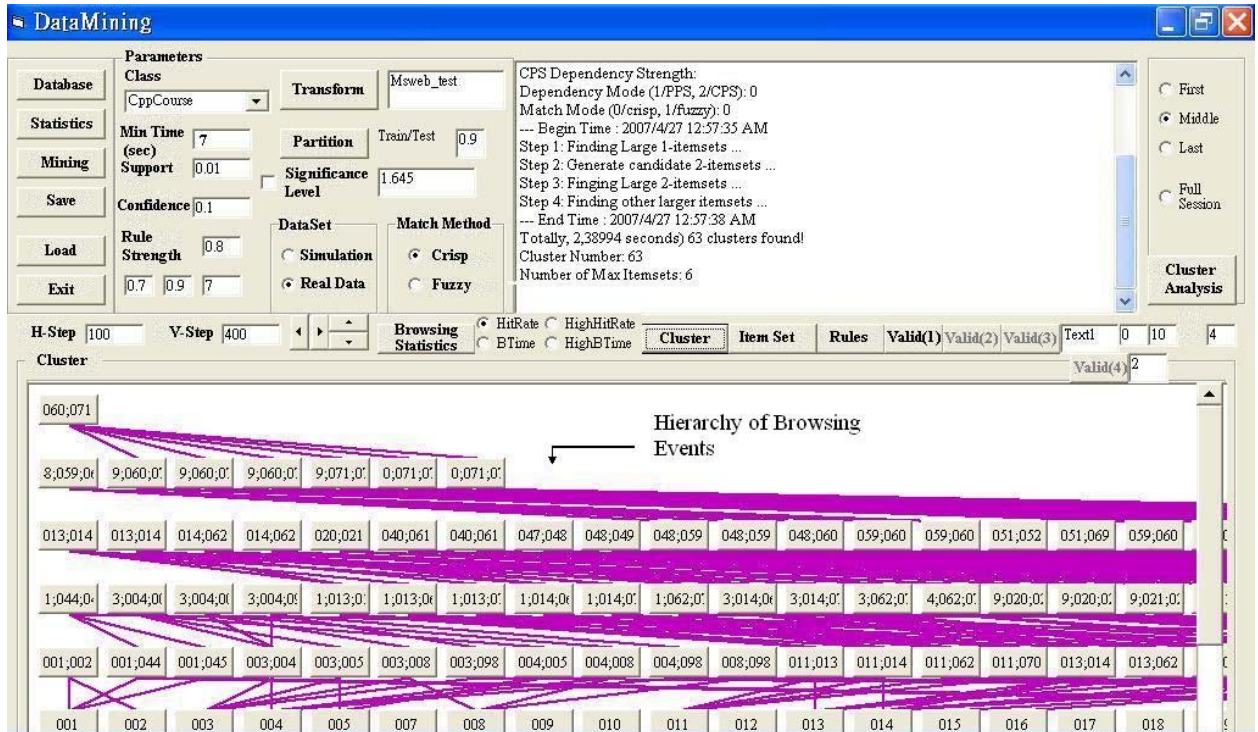


Figure 3: A screenshot (English-translated) of the association mining tool for discovering frequent browsing events and generating contextualized recommendation rules

As to the correlation between browsing events and learning achievement, it is computed by conducting a linear association test on the population correlation coefficient  $\rho$ ,  $-1 \leq \rho \leq 1$ , between the total browsing effort (TBE) and learning achievement of each student. The null hypothesis is  $H_0: \rho=0$ , which shows no linear correlation in the population, and the alternate hypothesis is  $H_a: \rho \neq 0$ , which means there exists a linear correlation (either positive or

negative) in the population. The test statistic is  $t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}}$ , where  $r$  is the Pearson correlation coefficient and  $n$  is the number of students.

#### The tool

We have implemented a mining tool, as shown in Figure 3, to discover frequent browsing events and generate recommendation rules based on the contextualized browsing model.

```

CPP950:
(?c browse fhwang*scoref*8629)
->
(?c sug_browse fhwang*scoref*8664) , (fhwang*scoref*8664 cf 0.81081*0.18072*0*1) |

CPP951:
(?c browse fhwang*scoref*8664)
->
(?c sug_browse fhwang*scoref*8629) , (fhwang*scoref*8629 cf 0.64516*0.18072*0*1) |

CPP952:
(?c browse fhwang*scoref*17416)
->
(?c sug_browse fhwang*scoref*8629) , (fhwang*scoref*8629 cf 0.4*0.01205*0*1) |

CPP953:
(?c browse fhwang*scoref*8629)
->
(?c sug_browse fhwang*scoref*8677) , (fhwang*scoref*8677 cf 0.7027*0.15663*0*1) |

CPP954:
(?c browse fhwang*scoref*8677)
->
(?c sug_browse fhwang*scoref*8629) , (fhwang*scoref*8629 cf 0.71233*0.15663*0*1) |

```

Figure 4: Some contextualized recommendation rules

The recommendation model is stored as a set of recommendation rules in Jena Triple format. Each rule contains at least one antecedent and a unique decision consequent. A variety of information, such as the author ID of learning objects, object type, object ID, and all the contextual information, are encoded and concatenated together with the separator character '\*', as shown below:

```

[RuleName:
  (?c browse author_id*object_type*object_id)
  (?c browse author_id*object_type*object_id)
  ....
  →
  (?c sug_browse author_id*object_type*object_id),
  (author_id*object_type*object_id cf confidence*support*group_relevance_type*achievement_relevance_type)],

```

where ?c denotes a variable representing some student.

For example, the following recommendation rule,  
[r0:

```
(?c browse fhwang*scoref*5197), (?c browse fhwang*scoref*6061), (?c browse fhwang*scoref*6056)
→
(?c sug_browse fhwang*scoref*6058), (fhwang*scoref*6058 cf 0.8*0.05545*0*0) ]
```

says that if a student has browsed the SCO object of ID=5197 (authored by ‘fhwang’), the SCO object of ID=6061 (authored by ‘fhwang’) and the SCO object of ID=6056 (authored by ‘fhwang’), then the rule will recommend the SCO object of ID=6058 (authored by ‘fhwang’) with confidence=0.8. By SCO here we mean Sharable Content Object, which is a concept derived from SCORM. Furthermore, the browsing event is group-independent and achievement-irrelevant, with support=0.05545 and confidence=0.8. Figure 4 lists some contextualized recommendation rules discovered by the mining tool.

## The Learning Environment

### The IDEAL system

IDEAL is a standard-conformant and Web-based learning content management system that was developed to help instructors design and deliver reusable learning objects conformant to modern eLearning standards like IEEE LOM, ADL’s SCORM, IMS QTI, and IMS RS. The motivation of developing IDEAL is to integrate three kinds of learning resources, including learning content, test items and learning design, conformant to currently popular e-Learning standards. In particular, IEEE LOM is the main metadata-tagging scheme adopted in IDEAL. Through the LOM metadata tagging, learning objects in IDEAL can carry educational contextual information that is helpful to their applications in personalized learning. Besides, IDEAL provides several educational facilities like course management, on-line testing and reporting services, resource access tracking, and statistical report of user feedbacks on learning objects. Through the IDEAL system, students can take on-line courses, access learning objects and engage themselves in on-line testing. Therefore, IDEAL can be deployed as a computer-assisted learning environment on the Web. Currently, IDEAL is running on Tomcat 5.0 with a native XML database X-Hive 6.0.

### Search Result List

Totally 16 records found

1—10 (Page 1/2)

#### [Private Resources]

1. XPATH Package (manifest)  
**Classification:** (/SimpleResourceType/Theory)(/SimpleResourceType/Exercise)  
**Description:** This package contains slides and an online exercise for the  
                   XPATH topic  
**Difficulty:** moderate  
**Producer (citer):** fhwang (fhwang@mcu.edu.tw)  
**Popularity:** unavailable
2. XPATH Sample Collections (manifest)  
**Classification:** (/SimpleResourceType/Example)  
**Description:** This learning objects contains a set of XPATH samples  
  
**Difficulty:** moderate  
**Producer (citer):** fhwang (fhwang@mcu.edu.tw)  
**Popularity:** unavailable

Figure 5: A screenshot (English-translated) for searching learning objects in the IDEAL system

In this research, IDEAL serves as a Web-based repository and a deliver of learning objects. IDEAL provides two channels for students to access learning objects. One is through the index list of learning packages, which are pre-

designed by instructors packing relevant learning objects together for specific topics. The second is through a search component in IDEAL. Figure 5 shows a screenshot of a list of learning objects returned by the search component in IDEAL. Specifically, the search component is a flexible search engine that could search learning resources at different levels of granularity. A search criterion in IDEAL is described by three categories of query constraints: (1) the LOM metadata constraint, (2) the document-space constraint indicating the resources created by specific authors, and (3) the access-control constraint. The LOM constraints are described in terms of IEEE LOM entries, such as the title, keyword, learning resource type and so on. The document-space constraint describes where the resources could be located, either in public or private user spaces. The last constraint describes the access control criterion, which in the current implementation says that students have to enroll in a course before they can access the resources authorized to the course.

### The recommender

A recommendation server was designed based on the Web-service technology to provide open recommendation services. An automatic reasoning system written in Jena performs the inference task based on the contextualized rules and the current browsing history of a student. The recommendation list is passed back to IDEAL in XML format, and then IDEAL transforms the recommendations from XML format into HTML pages using the XSL technique. The XSL technique was adopted to implement different recommendation modules. For each recommendation module, one XSL program is coded to display the recommendations showing different types of contextual information.

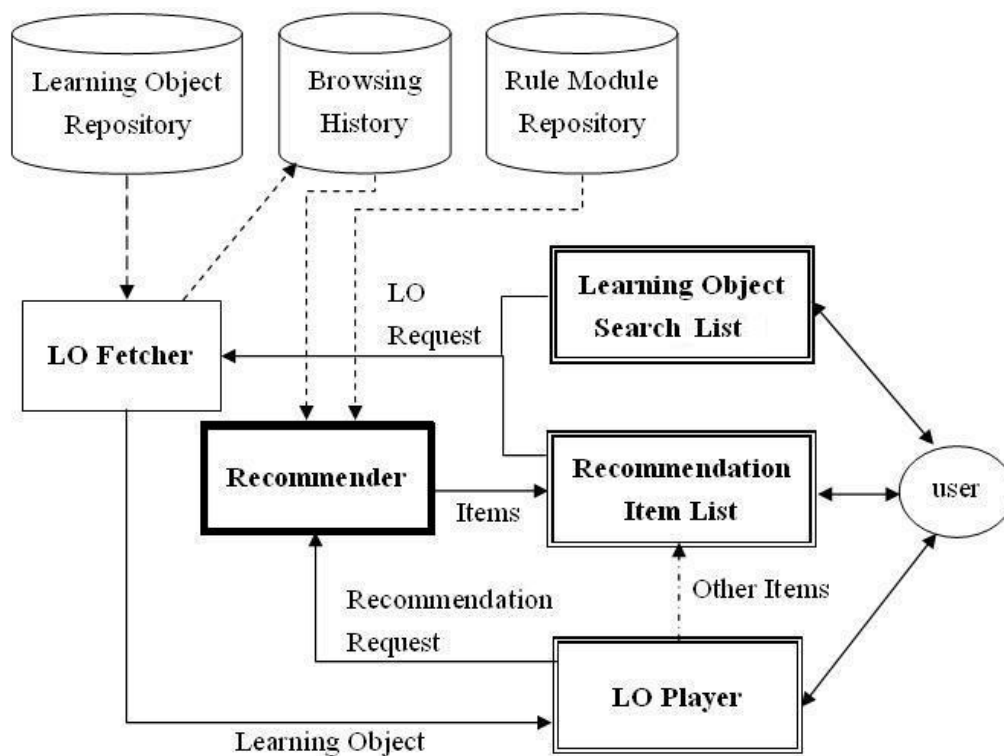


Figure 6: The interaction among the recommender, IDEAL and a user

The interaction among the recommender, IDEAL and a student is depicted as shown in Figure 6. A student first selects an item from the search list. This will trigger a learning object request (LO request) to the LO Fetcher in IDEAL, and the LO Fetcher will then retrieve the requested learning object from the LO repository and send it to the learning object player (LO Player). The LO Player is aimed to display the content of a learning object while providing interactions between the learning object and the student through several embedded buttons in the screen. The interactive buttons include those for posting/reviewing comments concerning the learning object, and those for



making recommendation requests. Figure 7 shows a screenshot of browsing a learning object in IDEAL, where students can post comments and/or browse comments given by other students about the learning object. Besides, students can ask for item recommendations from the two buttons in the top-right part of the screen.

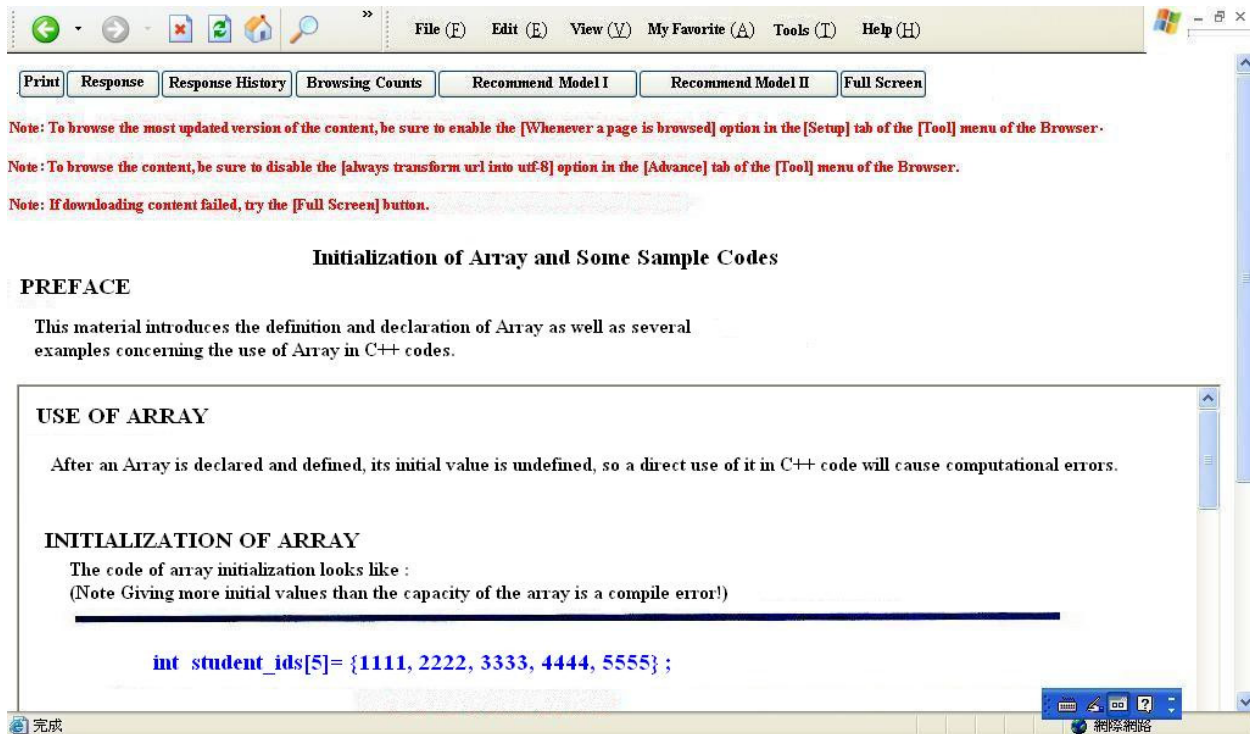


Figure 7 A screenshot (English-translated) for browsing learning objects in the IDEAL system

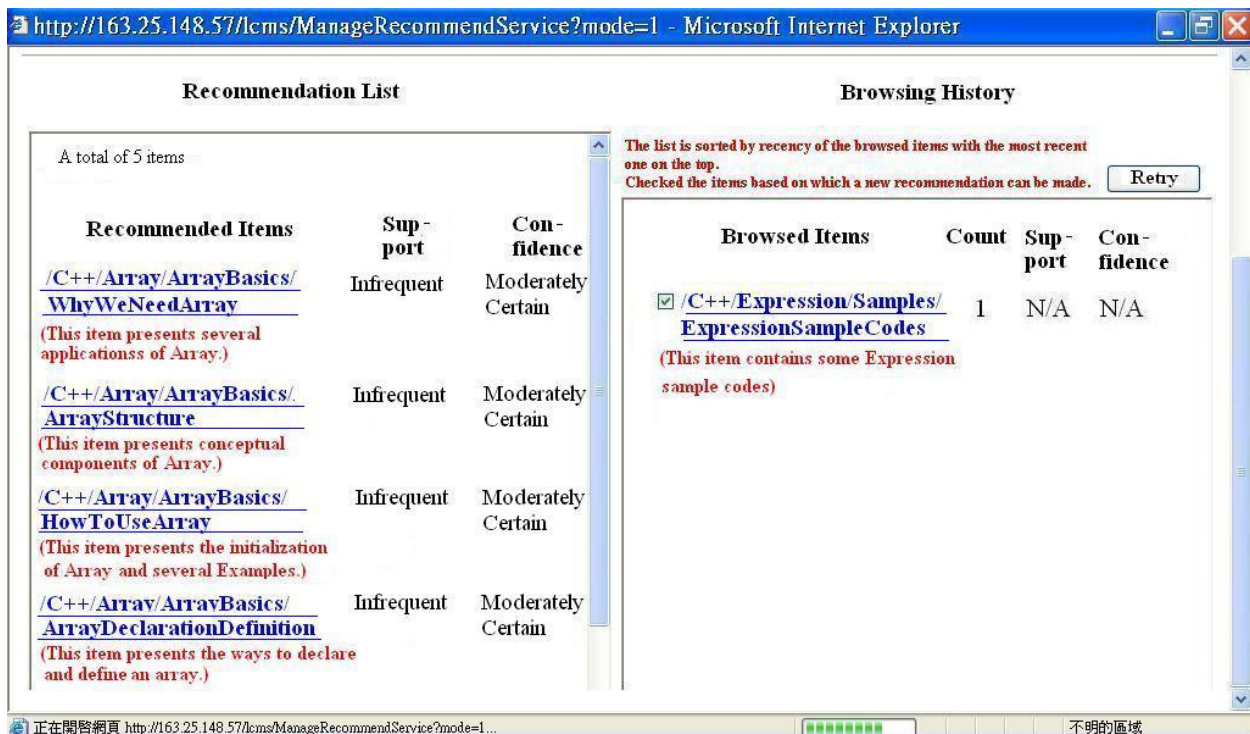


Figure 8: A screenshot (English-translated) for recommendation module of pure association mining showing only support and confidence in the IDEAL system

The recommender currently provides two recommendation modules. The first one displays only support, confidence and object description for each recommendation item. Figure 8 shows a screenshot of the recommendation list displayed by the first module. In the left window, the recommendations are given in a descendant order of confidence, which is classified into five categories ranged from “least-certain” to “very-certain”. Students can select recommendations based on support, confidence, and/or object descriptions. Furthermore, a student can ask for a totally new recommendation list by switching on/off the items listed in the browsing history window on the right, in which the items are sorted in a descendant order based on recency of the browsing time. This is especially helpful when students want the system to focus on more recent browsing events that reflect their most recent interests.

The second module displays recommendations according to the educationally contextualized event model. In addition to support, confidence and object descriptions, a variety of educational contextual information like group relevance and achievement relevance are also displayed. Figure 9 shows a screenshot of the recommendation list rendered by the second module. Students are provided with links associated with different types of recommended items. Therefore, they can make a decision on which recommendations to take based on their preference for contextual information. By selecting one of the links, the recommendation window below the Link one will display a list of recommended items. Again, students can ask for a totally new recommendation by switching on/off the items listed in the browsing history window on the right.

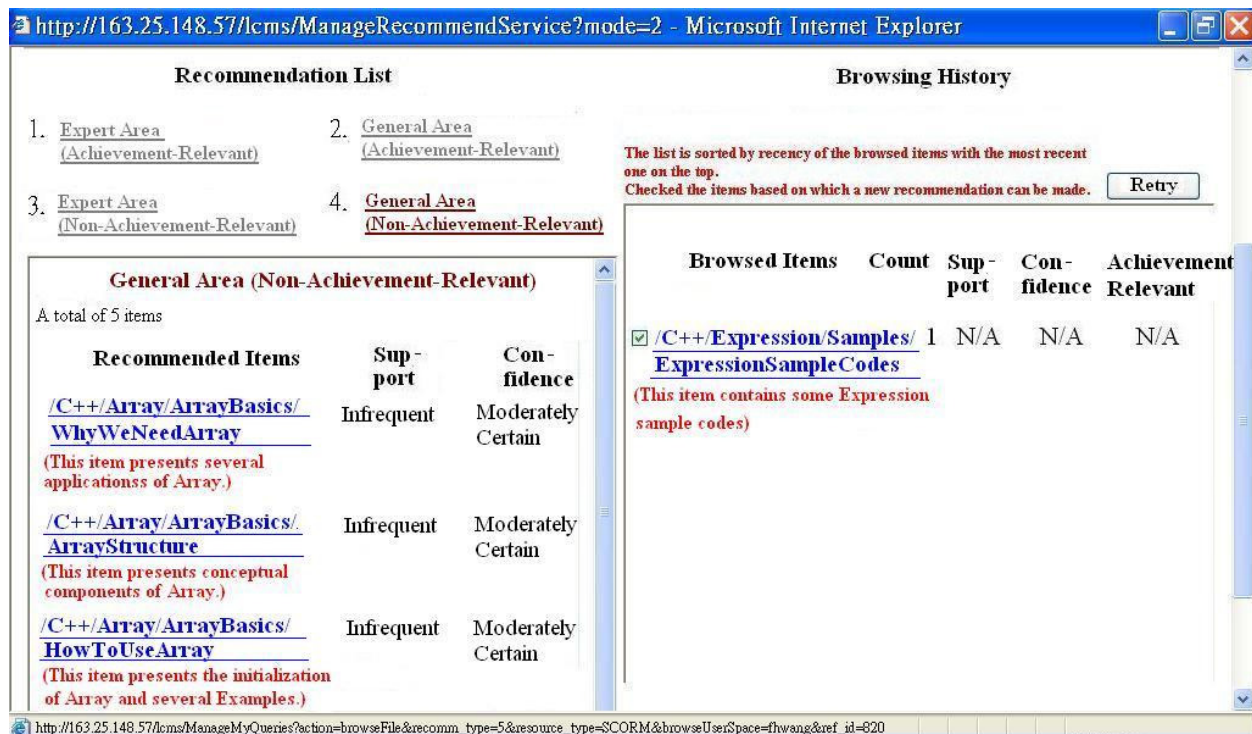


Figure 9: A screenshot (English-translated) for recommendation module of contextualized browsing model showing more educational contextual information in the IDEAL system

## Evaluations and Results

A prototype evaluation of the system by a few experts had been conducted so that feedback on the system design and quality of the recommendation services can be collected (Wang, 2007). Problems exposed through the feedback have been corrected. To further assess the revised system, another evaluation with a larger number of students was conducted. Furthermore, an evaluation was conducted to compare the performance of the recommender based on pure association mining with that of the one based on the contextualized browsing model. At last, suggestions and insights gained from the research into personalized learning are also presented.

## Model construction

To build a test recommendation model, a course named “XML Programming” was first conducted from February 2006 to July 2006. A total of 65 students enrolled in the course. Given the free network access to on-line resources in IDEAL, all fetched learning objects during each learning session are tracked and stored in a log database. For each learning activity in the course, the learning achievement of each student was measured by a post test and stored in a database for further analysis. Table 2 shows a summary of the mining results obtained from these access traces. As shown in Table 2, a total of 477 browsing events were discovered.

Table 3 lists the contextualization results over the 477 browsing events, with  $\text{support} \geq 0.02$ , under the  $F$  distribution with (2, 62) degrees of freedom, and the  $t$  distribution with 60 degrees of freedom, and  $P\text{-value} \leq 0.05$ . It can be seen that only 8.81% of the browsing events are group-relevant ones, all of which happened to be characteristic of high-achievement students. Though just only a small percentage, these events represent special browsing behavior of the high-achievement students. Besides, it is shown in Table 3 that 10.06% of the browsing events are achievement-relevant, all of which are positively relevant. Furthermore, it is interesting to find that none of the browsing events are negatively related to learning achievement.

Table 2: Summary of browsing events in a sample course

Summary	Course Name XML Course
Total Learning Activities	6
Total Log Records	10924
Total User Sessions	1082
Average Session Length	3.67
Number of Learning Objects	64
Minimum Support	0.02
Maximum Event Length	10
Number of Discovered Events	477

Table 3: Contextualization results of browsing events in a sample course

Group-relevance Achievement-correlation	Group-Irrelevant	Group-Relevant	Total
Positively Correlated	20	28	48 ( <b>10.06%</b> )
Non-Correlated	415	14	429 (89.94%)
Negatively Correlated	0	0	0 (0%)
Total	435 (91.19%)	42 ( <b>8.81%</b> )	477

## Student Evaluation

The recommender was then evaluated by another 36 undergraduate students who enrolled in the XML programming course in 2007 spring semester. Among the 36 students, four students claimed themselves as experts, and the other thirty-two students as novices. After a brief introduction to the system operation, these students were given three hours to study the course using the recommender system. After the self-study hours, they were asked to fill out a questionnaire, and a post test was administered to them. The questionnaire contains thirteen 7-point Likert scale items, the same as those used in the prototype evaluation (Wang, 2007). Table 4 shows the questionnaire evaluation results.

From Table 4, it can be seen that students were satisfied with the efficiency (with mean = 5.06) and accuracy (with mean = 5.08) of the system. Besides, as to the accuracy of the recommendations based on different types of contextual information, it is interesting to find that students were satisfied with the recommendations based on the information of achievement relevance (with mean = 5.08), group-relevance (with mean=5.0) and confidence (with mean=5.0). Finally, students agreed the system, as a whole, is beneficial to their learning (with mean =5.0).

As to the preference, sufficiency and usefulness of the contextual information, students felt moderately positive toward them. Since only moderately positive attitudes toward the aforementioned items are revealed, further investigation into the decision styles of different student groups will be made and discussed later.

The questionnaire also contains four open questions related to how students interact with the recommender. A summary of student answers to the four questions are listed in Table 5. The first question is about which type of information students considered the most important and the least important, respectively, for selecting recommendations. Table 5 shows that information of group independency is considered the most important by 22.2% students, while only 8.3% students considered it the least one. Furthermore, 13.9% students valued the information of confidence, while only 5.5% students disfavoring it. Therefore, information of confidence was the second welcome one. On the other hand, only 8.3% students considered the information of object description the most important, while 22.2% students considered it the least important. As some students had expressed in an interview that the object descriptions were a little bit rough, and hence they failed to provide useful information for making sensible decisions.

*Table 4: Results of student evaluation*

Question Item Description	Mean	Std. Dev.
1. The system is friendly to use. (friendliness)	4.72	1.23
2. The system is efficient in generating recommendations.(efficiency)	5.06	1.27
3. What I need can be found in the recommendation list. (recall)	4.72	0.97
4. Items in the recommendation list meet my needs. (accuracy)	5.08	0.94
5. I would like to use the various types of information to sort out what I need in the recommendation list. (preference)	4.78	1.07
6. The information is sufficient to help me sort out what I need in the recommendation list. (sufficiency)	4.58	1.00
7. The information is helpful in sorting out what I need in the recommendation list. (usefulness)	4.89	0.92
8. The recommendations given by group-dependent browsing events meet my needs.	5.00	0.99
9. The recommendations given by group-independent browsing events meet my needs.	4.53	0.84
10. The recommendations given by support meet my needs.	4.94	1.07
11. The recommendations given by confidence meet my needs.	5.00	0.93
12. The recommendations given by achievement-relevance meet my needs.	5.08	0.97
13. The contents recommended by the system are helpful to my learning. (effectiveness)	5.00	1.04

*Table 5: Some answers to the open questions of the questionnaire in student evaluation*

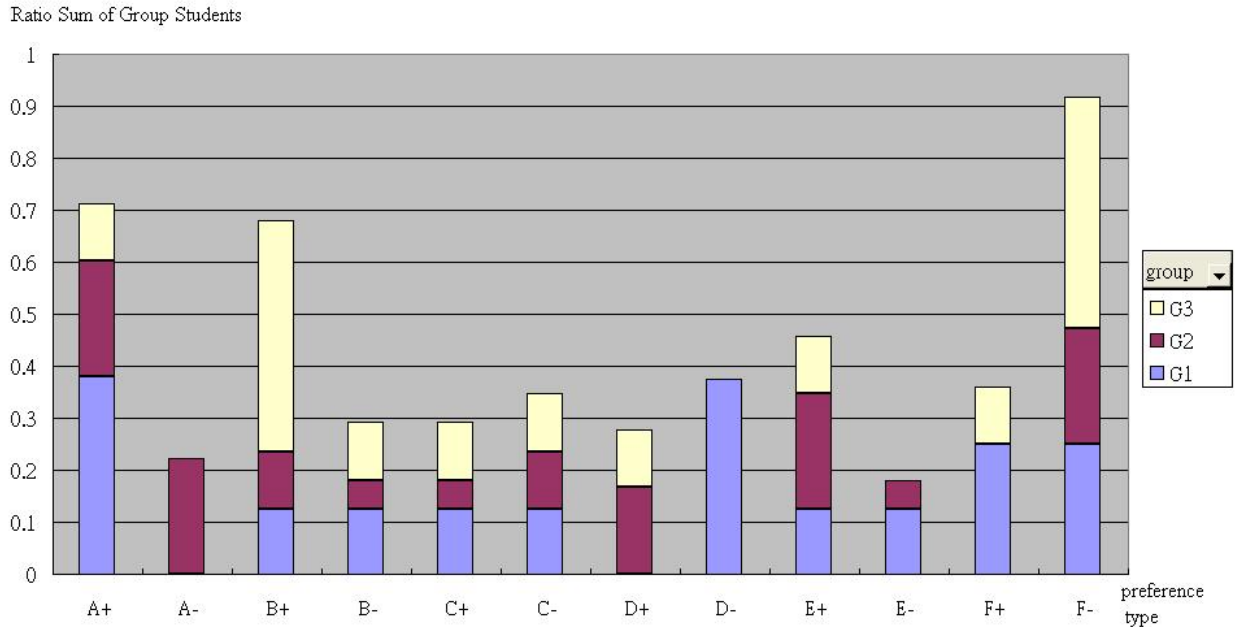
Question	Answers		
1. Which information do you think is the most important for selecting recommended learning resources? Which is the least one?	Information	Most Important	Least Important
	1. Group Relevance	7 (19.4%)	6 (16.6%)
	2. Group Independency	8 (22.2%)	3 (8.3%)
	3. Learning Achievement Relevance	4 (11.1%)	5 (13.9%)
	4. Support	4 (11.1%)	3 (8.3%)
	5. Confidence	5 (13.9%)	2 (5.5%)
	6. Object Description	3 (8.3%)	8 (22.2%)
2. What other types of information you think are needed to support the recommendation decision?	2-1. Average time spent for browsing a learning object 2-2. A significant index provided by peer students 2-3. Recommendations provided by high-achievement students 2-4. Recommendations for novices only		
3. Do you have any suggestions for improving the system interface?	3-1. The interface design is too complex (too many sub-windows). 3-2. Relate the recommendations based on some semantic meanings.		
4. Which recommendation model do you prefer?	4-1. Twenty two vote for pure association module (61.1%) 4-2. Thirteen vote for contextualization module (36.1%)		



Roughly equal percentage of students favored and disfavored each of the remaining three types of information (group relevance, achievement relevance, and support). They seemed to be the most controversial ones. In particular, relatively high percentage students favor (with 19.4%) and disfavor (with 16.6%) the information of group relevance, respectively. Similar situations happened to the information of achievement relevance. It seemed that the students had shown a huge controversy over the preference for educational contextual information (group relevance and achievement relevance).

### Group comparison and results

Students are divided into three groups according to their post test scores. The high-achievement group consists of the top 25% students, the middle-achievement group consists of the next 50% students, and the remaining 25% students are in the low-achievement group. To investigate further into the information preference specific to each student group, the ratios of students favoring and disfavoring the various information types for each student group are depicted in Figure 10, where G1 denotes the high-achievement group, G2 denotes the middle-achievement group, and G3 the low-achievement group.



*Figure 10:* Ratios of students favoring and disfavoring the information types for each student group (A+: favor Group Relevance; A-: disfavor Group Relevance; B+: favor Group Independency; B-: disfavor Group Independency; C+: favor Achievement Relevance; C-: disfavor Achievement Relevance; D+: favor Support; D-: disfavor Support; E+: favor Confidence; E-: disfavor Confidence; F+: favor Object Description; F-: disfavor Object Description)

For high-achievement students, they favored the group-relevance information most. Therefore, a large number of high-achievement students seemed to be highly motivated to visit what past high-achievement students had browsed. Besides, they seemed to have the profession to comprehend the short and concise object descriptions in order to select sensible recommendations. On the other hand, near 30% of high-achievement students disfavored the information of support, implying the tendency to ignore the frequency of past browsing events.

For middle-achievement students, 22% students favored the information of group relevance; while there were also near 22% students disfavored that. Besides, another 22% of middle-achievement students favored the information of confidence and the same percentage of students disfavored the information of object descriptions. Therefore, middle-achievement students seemed not to be as motivated as high-achievement students in following the behavior of past high-achievement students. Nor did they have the same profession as high-achievement students to comprehend the concise object descriptions. Besides, they seemed to rely more on the information of confidence to select sensible recommendations.

For low-achievement students, 44% of them favored the group independency. On the other hand, there were near 44% students disfavored the information of object descriptions. Therefore, for low-achievement students, they are the least motivated to follow past high-achievement students, and have the least profession to comprehend the concise object descriptions.

## Performance Evaluation

To evaluate the performance of the proposed recommendation model, 80 students from two “C Programming” courses joined the experiment. Among them, 40 students were randomly picked out and assigned to the control group, who used the recommender with the pure association mining model. The other 40 students were assigned to the experiment group, who used the recommender with the contextualized browsing model. Before the experiment started, all students took a pretest on the topics of “variables”, “flow-control” and “recursive programming”. After a brief introduction to the system operation, they were given three hours to study those topics. After the self-study hours, a post test was administered to them. Table 6 shows the results of both tests. The *t*-test was used to test the difference between the score means of the two experiment groups. As shown in Table 6, there is no significant difference between the pretest mean scores of both student groups. After the treatment was administered, it can be seen that the mean score of the experiment group was significantly higher than that of the control group.

Table 6: Test results in the performance experiment using *t*-Test of 77 df

Group \ Test	Pretest		Post-Test	
	Mean	Std.	Mean	Std.
Control Group	56.95	15.31	45.83	16.14
Experiment Group	59.30	14.83	53.03	15.33
<i>P</i> -value	0.244		0.013*	

\* Using significance level of 0.05.

## Research suggestions

As a summary, the main contributions of this research can be described from three perspectives. The first is concerned with how students react to the various types of contextual information. The second is concerning how well students think the system has done the jobs. Finally, the last is related to the performance of the new recommendation model. In this research, six types of contextual information were investigated, and three of them (group relevance, achievement relevance, object descriptions) are educational ones, and the other three (group independence, support, confidence) are non-educational.

### *Perspective 1: How students react to the various types of contextual information*

As a whole, students react differently for different information types. Specifically, high-achievement students are more willing to follow the behavior of past high-achievement students, while low-achievement students show the least preference. Middle-achievement students seemed not to be as motivated as high-achievement students in following the behavior of past high-achievement students. Besides, they seemed to rely more on the information of confidence to make sensible decisions on adopting which recommendations. The difference in preference may also come from the different needs of learning control specific to different student groups. Low-achievement students are less willing to take on extra responsibility to make the learning decision, while high-achievement students tend to learn autonomously.

Therefore, as low-achievement students heavily rely on direct supports from the system, the recommendations given to them should be as useful and simple as possible (e.g., a complete package of learning objects). The complex decision model can therefore be hidden in the background from them. For high-achievement students, educational contextual information does help, especially the information of group relevance and object descriptions. Furthermore, they tend to ignore the non-educational information of support of past browsing events. Finally, for middle-achievement students, a controversy exists among the preferences for various information types. As a result,

there is a need of providing the support from multiple perspectives, including educational and non-educational ones, for middle-achievement students to make sensible decisions on taking recommendations.

#### *Perspective 2: How well students think the system has done the jobs*

As a whole, the efficiency of the recommender is acceptable to most students. However, the friendliness of the system is simply moderately acceptable. As aforementioned, different student groups have different needs for the decision model; therefore, the system interface should also be adaptable to different student groups. Furthermore, the accuracy of recommendation is more acceptable than the recall rate, especially for recommendations generated based on the educational contextual information.

#### *Perspective 3: The performance of the contextualized browsing recommendation model*

The preliminary experiment showed that students using the recommender with contextualized browsing model gained higher achievement than those using the recommendation model with pure association mining. It showed that contextualizing browsing events is a potential way to improve the performance of the recommender.

### **Concluding Remarks**

In this paper, a new method for discovering contextualized browsing events based on association mining and statistical techniques is proposed. Also presented is a Web-based LCMS system, called IDEAL, which fulfills the idea of designing and disseminating personalized learning objects. As outlined in Perspective 2 of the previous section, the current data mining approach, though providing useful browsing experiences, fails to take care of each student's individual needs. Therefore, more types of contextual information, such as the learning style, learning orientation (Martinez, 2001), and student learning states, could also be considered in order to mine out more personal learning patterns. As a result, it is worth acquiring the knowledge of more types of educational contextual information so that their effects on personalized learning can be investigated in future research.

### **Acknowledgement**

The author would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 95-2520-S-130 -001. The author would like to express sincere thanks to the anonymous reviewers for their valuable comments on this paper.

### **References**

- Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, 22(2), 207–216.
- Agrawal, R., & Srikant, R. (1994). Fast algorithm for mining association rules. In Bocca, J. B., Jarke, M. & Zaniolo, C. (Eds) *Proceedings of 20th International Conference on Very Large Data Bases (VLDB'94)*, Santiago de Chile, Chile, Morgan Kaufmann, 487-499.
- Andris, J. (1996). The relationship of indices of student navigational patterns in a hypermedia geology lab simulation to two measures of learning style. *Journal of Educational Multimedia and Hypermedia*, 5, 303–315.
- Aroyo, L., & Dicheva, D. (2001). Aims: Learning and teaching support for www-based education. *International Journal for Continuing Engineering Education and Life-long Learning*, 11, 152–164.
- Brennan, M., Funke, S., & Anderson, C. (2001). The Learning Content Management System: A New e-Learning Market Segment Emerges. *IDC White paper*, retrieved July 10, 2008 from <http://www.internetttime.com/Learning/lcms/IDCLCMSWhitePaper.pdf>.

- Brennan, M., Funke, S., & Anderson, C. (2001). The Learning Content Management System: A New e-Learning Market Segment Emerges. *IDC White paper*, retrieved July 10, 2008 from <http://www.internetttime.com/Learning/lcms/IDCLCMSWhitePaper.pdf>
- Brusilovsky, P., & Vassileva, J. (2003). Course sequencing techniques for large-scale Web-based education. *International Journal for Continuing Engineering Education and Lifelong Learning*, 13, 75-94.
- Calcaterra, A., Antonietti, A. & Underwood, J. (2005). Cognitive style, hypermedia navigation and learning. *Computers and Education*, 44, 441-457.
- Chen, S. Y., & Macredie, R. D. (2002). Cognitive styles and hypermedia navigation: development of a learning model. *Journal of the American Society for Information Science and Technology*, 53(1), 3-15.
- Chen, M. S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866-883.
- Denaux, R., Dimitrova, V., & Aroyo, L. (2004). Interactive ontology-based user modeling for personalized learning content management. In de Bra, P. and Nejdl, W. (Eds) *Proceedings of the SWEL Workshop at Adaptive Hypermedia, Lecture Notes in Computer Science* 3137, Springer Verlag, 338-347.
- Ford, N., & Chen, S.Y. (2000). Individual differences, hypermedia navigation and learning: An empirical study. *Journal of Educational Multi-media and Hypermedia*, 9(4), 281-312.
- Kinshuk, & Lin, T. (2003). User exploration based adaptation in adaptive learning systems. *Journal of Information Systems in Education*, 1(1), 22-31.
- Liang, T., & Leifer, L. J. (2000). Re-use or re-invent? Understanding and supporting learning from experience of peers in a product development community. In Budny, D. & Bjedov, G. (Eds.) *Proceedings of the 30th ASEE/IEEE Frontiers in Education Conference*, Kansas City, MO, USA, Champaign: Stipes Publishing L.L.C, S3A 14 - S3A 19.
- Liu, M., & Reed, W. M. (1994). The relationship between the learning strategies and learning styles in a hypermedia environment. *Computers in Human Behavior*, 10, 419-434.
- Liu, M., & Reed, W. M. (1995). The effect of hypermedia assisted instruction on second-language learning through a semantic-network-based approach. *Journal of Educational Computing Research*, 12(2), 159-175.
- Martinez, M. (20001). Key Design Considerations for Personalized Learning on the Web. *Educational Technology & Society*, 4(1), 26-40.
- McNaught, C., Kennedy, D., & Majoor, J. (2002). Designing online learning sites to cater for learning needs. In Kinshuk, Lewis, R., Akahori, K., Kemp, R., Okamoto, T., Henderson, L. Lee, C.-H. (Eds.) *Proceedings of International Conference on Computers in Education*, Auckland, New Zealand, IEEE Computer Society Press, 535-539.
- Mittal, A., Krishnan, P. V., & Altman, E. (2006). Content classification and context-based retrieval system for eLearning. *Educational Technology & Society*, 9 (1), 349-358.
- Mobasher, B., Cooley, R., & Srivastava, J. (2000). Automatic personalization based on web usage mining. *Communications of the ACM*, 43, 142-151.
- Mohan, P. (2004). Reusable on-line learning resources: problems, solutions and opportunities. In Kinshuk, Looi, C-K, Sutinen, E., Sampson, D., Aedo, I., Uden, L. & Kähkönen, E (Eds.), *Proceedings of IEEE International Conference on Advanced Learning Technologies*, Los Alamitos, CA: IEEE Computer Society, 904-905.
- Mohan, P., & Greer, J. (2003). Reusable learning objects: current status and future directions. In Lassner, D. & McNaught, C. (Eds.), *Proceedings of ED-MEDIA 2003 World Conference on Educational Multimedia, Hypermedia and Telecommunication*, Honolulu, Hawaii, USA, Chesapeake, VA: AACE, 257-264.
- Mulvenna, M. D., Anand, S. S., & Buchner, A. G. (2000). Personalization on the net using Web mining. *Communications of the ACM*, 43(8), 123-125.
- Najjar, J., Wolpers, M., & Duval, E. (2006). Towards effective usage-based learning applications: track and learn from user experience(s). In Kinshuk, Koper, R., Kommers, P., Kirschner, P., Sampson, D. & Didden, W. (Eds), *Proceedings of IEEE International Conference on Advanced Learning Technologies*, Los Alamitos, CA: IEEE Computer Society, 1022-1024.

- Papanikolaou, K.A., & Grigoriadou, M. (2003). An instructional framework supporting personalized learning on the Web. In Devedzic, V., Spector, J. M., Sampson, D. G. & Kinshuk (Eds.), *Proceedings of IEEE International Conference on Advanced Learning Technologies*, Los Alamitos, CA: IEEE Computer Society, 120 – 124.
- Parkinson, A., & Redmond, J. A. (2002). Do cognitive styles affect learning performance in different computer media? *ACM SIGCSE Bulletin*, 34, 39-43.
- Sampson, D., Karagiannidis, C., & Cardinali, F. (2002). An architecture for Web-based e-learning promoting reusable adaptive educational e-content. *Educational Technology and Society*, 5(4), 27-37.
- Siegel, S., & Castellan Jr., N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences* (2<sup>nd</sup> Ed.), New York: McGraw-Hill.
- Wang, F.-H., & Shao, S. M. (2004). Effective personalized recommendation based on time-framed navigation clustering and association mining. *Expert Systems with Applications*, 27(3), 365-377.
- Wang, F.-H. (2007). A Contextualization Method of Browsing Events in Web-based Learning Content Management System for Personalized Learning. In Spector, J. M., Sampson, D. G., Okamoto, T., Kinshuk, Cerri, S. A., Ueno, M. & Kashihara, A. (Eds.), *Proceedings of IEEE International Conference on Advanced Learning Technologies*, Los Alamitos, CA: IEEE Computer Society, 43–45.
- Yin, L. R. (2001). Dynamic learning patterns: Temporal characteristics demonstrated by the student. *Journal of Educational Multimedia and Hypermedia*, 10, 273–284.
- Zaiane, O. R. (2001). Web usage mining for better Web-based learning environment. In Hartley, J. R., Okamoto, T., Kinshuk & Klus, J. P. (Eds.), *Proceedings of IEEE International Conference on Advanced Learning Technologies*, Los Alamitos, CA: IEEE Computer Society, 60-64.

## Using Teradata University Network (TUN), a Free Internet Resource for Teaching and Learning

**Robert Winter, Anke Gericke and Tobias Bucher**

Institute of Information Management, University of St. Gallen, Switzerland // robert.winter@unisg.ch // anke.gericke@unisg.ch // tobias.bucher@unisg.ch // Tel.: +41 71 224 2190 // Fax: +41 71 224 2189

### ABSTRACT

Business intelligence and information logistics have become an important part of teaching curricula in recent years due to the increased demand for adequately trained graduates. Since these fields are characterized by a high amount of software and methodology innovations, teaching materials and teaching aids require constant updating. Teradata has teamed up with lecturers and researchers to build and run a portal to support teaching business intelligence and information logistics. This article describes how faculty can use the Teradata University Network (TUN) to prepare and run courses by reusing teaching materials and running state-of-the-art commercial software provided in an application service provider model. It furthermore describes experiences with an actual course on management information systems taught by the authors. Students' feedback on the course design is presented and conclusions are drawn on two similar courses. Our action research results show that students' adoption of and feedback about such courses has been very encouraging.

### Keywords

Business intelligence, Information logistics, Data warehousing, Teaching resources, Teaching platform

### Introduction

Business intelligence comprises of a broad range of business-oriented topics (e.g. customer relationship management, business performance management) as well as technical topics (e.g. data warehousing, data mining, large databases). Due to its significance for business success, the demand for software in these fields as well as for the underlying information logistics infrastructure has increased during the last years and the strong demand will continue (Wixom & Watson, 2001). The increasing importance of business intelligence is supported by the forecasted growth rate of the business intelligence market as well as a recent Gartner survey which reports that “business intelligence jumped from the No. 10 slot to the second-highest priority on CIOs' agendas” (Graham, 2005). The increasing demand for business intelligence solutions is also linked to a demand for qualified personnel who are able to use business intelligence solutions and develop or manage the underlying information logistics infrastructure. For that reason it is necessary that academics include these topics in their curricula and impart this knowledge to their students. However it is not sufficient that students master the theoretical fundamentals: From a company's point of view, it is much more important that their future employees can handle real-world problems—instead of small textbook examples—and master state-of-the-art commercial tools—instead of research prototypes. That is why courses in the field of business intelligence should not only comprise of theoretical contributions and fundamental concepts, but should also enable students to use and understand current software tools.

The Teradata University Network (TUN) initiative addresses this issue. It was developed in cooperation with Teradata by a group of scholars in 2001. The Teradata University Network is a portal which can be used free of charge by lecturers and students in the fields of data warehousing, business intelligence/management information systems (MIS), and database. Academia is supported not only by providing a teaching material exchange, but in particular by providing easy and free access to centrally operated and state-of-the-art commercial software tools in these fields. The access to these tools is implemented as an application service provider model (Watson & Hoffer, 2003; Winter, 2004).

TUN comprises a total user group of almost 1,800 registered faculty members from 890 universities and colleges in 72 countries. Approximately 69% of the registered faculty members are located in North America, South America, Central America and the Caribbean, 13% are located in Europe, the Middle East and Africa, and 13% are located in the Asia Pacific Region. Within the last twelve months, the Teradata University Network was actively used (e.g. by viewing pages and/or contributing content) by almost 700 faculty members (all figures as of 2007-11-30).

In the next section, we position the Teradata University Network within the range of existing online resources for learning/teaching support in the field. An overview of goals and content of the Teradata University Network is presented. The section “TUN Access and Content Integration” provides information regarding access to and integration of new material into TUN. Different forms of using the Teradata University Network in academic teaching will be explained. The section “Application Case” illustrates how we used TUN to prepare and run courses in two master programs. We also report on students’ feedback and other lessons learned in that section. Finally, we identify the limitations of the Teradata University Network. A short summary is presented and suggestions for extensions are made in the final section “Conclusion and Outlook”.

## **Related Work**

TUN is a web-based platform that provides access to teaching materials and software tools to lecturers and students. However, the Teradata University Network is not an e-learning platform and can also not be compared to the so called “academy” or “university” programs from software vendors. The following is a comparison of the differences between e-learning platforms/software vendor programs and TUN. Afterwards, the Teradata University Network will be compared to similar platforms.

E-learning platforms make learning content available over the internet. They allow students to help support their learning process e.g. by presenting the contents already studied, presenting test results, providing communications tools, etc. (Baumgartner et al., 2005). Contrary to an e-learning platform, the Teradata University Network does not support the learning process directly. Lecturers can rather use TUN material to prepare in-class courses as well as e-learning/distance learning courses.

Next to e-learning platforms, various offers from software vendors exist under numerous labels (e.g. Oracle University—<http://education.oracle.com>, Microsoft IT Academy Program—<http://www.microsoft.com/education/msitacademy/>, Red Hat Academy—<https://www.redhat.com/training/academy/>) which provide documents and software tools to students and/or lecturers. Most of these services, however, are not free of charge and comprise of materials only from a single software vendor. In addition, software tools are usually not provided in an application service provider model so that lecturers and/or students have to install and manage such tools. In contrast, the Teradata University Network not only provides tools from several vendors that are partially integrated (e.g. by using the same data set), but also provides such tools in an application service provider model and free of charge.

Like the Teradata University Network, there are other portals that neither provide e-learning support nor represent a specific sales/support channel for software vendors. The Edna Portal (<http://www.edna.edu.au>), the Gateway to Educational Material (<http://www.thegateway.org>), and the platform “Multimedia Educational Resource for Learning and Online Teaching” (<http://www.merlot.org>) solely provide links to teaching material. In contrast, platforms such as EducaNext (<http://www.educanext.org>), Universal Brokerage Platform (<http://nm.wu-wien.ac.at/universal/>), and Ariadne (<http://www.ariadne-eu.de>, based on the Universal Brokerage Platform) store their teaching materials in a database—like TUN does. However, unlike the Teradata University Network, these platforms do not concentrate on a specific subject. They neither try to establish a common understanding in a certain field nor offer a common, keyword based access. Such platforms do not provide any software in an application service provider model. The Gateway to Educational Material offers links to software tools as well, but only freeware tools are included and these have to be installed on the lecturer’s/student’s computer.

## **Goals and Content of TUN**

TUN (<http://teradatauniversitynetwork.com>) has been created to:

- be a prime resource for knowledge about data warehousing, decision support systems, business intelligence, and databases,
- build an international community whose members share their ideas, experiences, and resources with others,
- serve as a bridge between academia and the world of practice.

Plans to create TUN were first announced in late 2001. In early 2002, the design was completed and testing began. After adding content, the Teradata University Network was promoted at different conferences and on web sites in summer 2002. The formal roll out in fall 2002 was followed by adding the Teradata database software, a dimensional modeling tool (research prototype), and MicroStrategy software through a web-based application service provider arrangement in 2003. In 2004 and again in November 2005, TUN moved to a new platform. By these migrations, the platform's layout was optimized, and content was made accessible by multiple indexing.

The teaching material provided by the Teradata University Network and the available software will now be explained, followed by a description of further TUN resources. TUN structures its content in two different ways: On the one hand, different teaching materials (e.g. articles, research reports, assignments, etc.) can be accessed according to subject matter (course type). That means a user can select material that focuses on data warehousing, business intelligence, or database. On the other hand, teaching material can be accessed according to the content type, which means that all available articles, research reports or assignments are presented to the user irrespective of the subject area.

The following teaching materials are available in the Teradata University Network (structured according to the content type as of 2007-06-06; *Figure 1* shows the structure according to both content type and course type):

- more than 80 articles,
- 18 assignments—including teaching notes,
- one book chapter,
- eight case studies,
- one link to a course web site,
- eight descriptions of the integration of different types of material (“integrated material”),
- eight podcasts,
- six PowerPoint presentations—including speaker's notes,
- one software project,
- nine research reports,
- 14 course syllabi—most at graduate level,
- seven tutorials,
- 26 web seminars.

Under the category “integrated material” descriptions are given to explain the use of different kinds of material in an integrated form for course preparation. To give an example, descriptions of teaching materials from data warehousing, business intelligence, and database can be integrated to support a “customer relationship management” teaching module.

Besides the valuable range of materials for course preparation, the most important feature of the Teradata University Network is that data warehousing software, business intelligence software, and database software as well as other tools are made accessible through an application service provider model. Colleges and universities do not have to select, obtain, install, and maintain any software. Instead, all included tools can be used free of charge, using just browsers and internet access. Not only is the software made available, but the Teradata database is already loaded with various data sets, including those from popular textbooks. A substantial tutorial basis can be accessed with both the Teradata database and the Application Modules of MicroStrategy. As of 2007-06-14, the following software tools and data sets have been available on TUN:

- Teradata database: Access to data sets from leading textbooks, including e.g.
  - the “Mountain View Community Hospital” case from the Jeffrey Hoffer et al. textbook “Modern Database Management” (Hoffer et al., 2004),
  - the database used in chapters 7 and 8 of “Modern Database Management” (Hoffer et al., 2004), and
  - the database used in Richard T. Watson's textbook “Data Management: Databases and Organizations” (Watson, 2003).
- Hyperion: TUN members are provided with free access to a business intelligence tool (Hyperion System 9 BI+) and a financial modeling tool (Hyperion System 9 Strategic Finance). The software is hosted by a different application service provider (CRESH.NET) and a separate registration is required to gain access to the Hyperion software.



- MicroStrategy 8: Since the beginning of 2007, the Teradata University Network has provided access to MicroStrategy 8. As of 2007-06-14, versions 7 and 8 ran in parallel but version 7 was taken out of service in fall 2007. The following modules are provided by MicroStrategy 8:
  - Customer analysis: e.g. customer segments analysis, analysis of the value proposition and the loyalty of customers, identification of cross selling possibilities, etc.,
  - Financial reporting analysis: analysis of receivables and payables, drawing the balance and profit and loss account, execution of future prospects, etc.,
  - Sales force analysis: analysis of product, distribution, and sales figures,
  - Sales and distribution analysis: analyzing the various factors that drive the sales and distribution process.
- University of Arkansas Resources: Access to large-scale, real-world datasets:
  - Sam's Club Database: A sales database provided by Wal-Mart Stores, Inc. containing six tables with more than 55 million rows.
  - Dillard's Department Stores Database: A sales database provided by Dillard's Inc. containing five tables and more than 128 million rows of retail sales.
  - Frozen Foods, Inc. Database: The data contains customer transactions for more than 13 thousand products shipped to 92 different sales districts in the United States. The database consists of six dimension tables that are linked to a fact table.
- Downloadable: Other easy-to-use, downloadable software e.g. a dimensional modeling tool and expert system software. In contrast to the main software offer, these software tools are research prototypes or evaluation copies with a restricted usage period.
- Demonstration: Various software demonstrations (e.g. for data quality management and data integration as well as a shell for expert systems).

TUN also comprises various resources from the commercial software training field. These resources however have not been peer reviewed:

- access to Teradata certification,
- over 100 white papers,
- over 150 web-based courses (1-2 hours each),
- Teradata user group,
- Teradata discussion forum,
- Tech Center: access to technical papers.

Besides the teaching materials and software tools listed above (see as well Figure 1), the mission of the portal is explained and the members of the TUN Advisory Board are presented. There is also an easy-to-use content submission form for new content or related web pages, and the latest newsletters as well as some quick start guides are provided. In order to gain fast access to the content of TUN, search functionality for document metadata was implemented in November 2005.

A separate portal for students, Teradata Student Network (<http://www.TeradataStudentNetwork.com>), provides access to a subset of the materials mentioned. Students are granted access to software tools, book chapters, articles, research reports, cases, projects, tutorials, podcasts and assignments—but not to syllabi, teaching notes and assignment solutions. Students can also use the web seminars and the tech center. By the end of November 2007, about 4,800 students from 395 universities and colleges in 62 countries have been registered for Teradata Student Network resources (all figures as of 2007-11-30).

## **TUN Access and Content Integration**

To gain access to the Teradata University Network, faculty must first register at the TUN web site. In addition to providing personal information such as name and university, applicants must provide a web page link that proves their teaching faculty status. This procedure is necessary because the material provided by TUN is—on the basis of fair use rules—only available to faculty members. Once their application has been authenticated (usually within 24 hours), faculty have access to the Teradata University Network. A separate registration is necessary for Teradata's SQL Assistant software which is coordinated with the Teradata University Network authorization. Registered TUN users can make the Teradata Student Network available to their students. A registration procedure for the Teradata

Student Network has been implemented in order to control database access (data sets are available for read-write access) and to collect usage metadata.

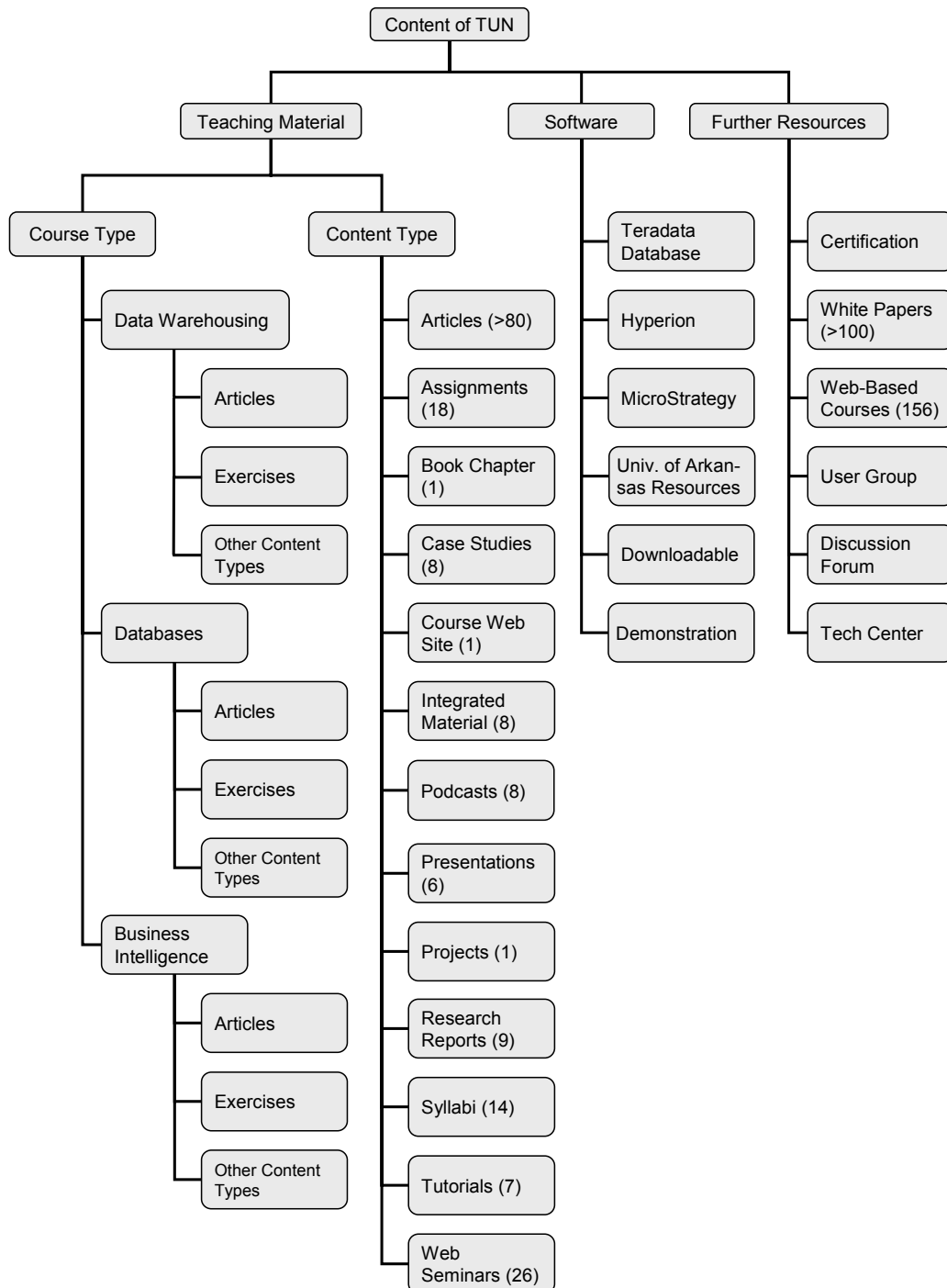


Figure 1. TUN content structure (all figures as of 2007-06-06)

What differentiates Teradata's initiative from other software vendors' resources for education is that leading academics are primarily responsible for the TUN vision, the evolution of TUN, and peer reviewing of submitted material. The Senior and Associate Directors work together with Teradata and other Advisory Board members as a management team in order to ensure that the Teradata University Network meets the needs of the IS academic

community. Project teams, led by board members, work closely with Teradata staff to make decisions, test prototypes, obtain beta testers, and make design decisions.

All content is reviewed by Senior and Associate Directors of the Board, i.e. by fellow faculty. As the value of the Teradata University Network increases with any further material, faculty are invited to submit material to the Director of the Board. There is an easy-to-use content submission form that collects the submitted metadata and allows the submitting faculty member to attach the content or provide a link for assessment. If appropriate, the reviewed content is released to the Teradata University Network and/or the Teradata Student Network.

## **Using TUN Resources to Support Teaching**

By using TUN resources, diverse instruction forms in the fields of data warehousing, business intelligence and database can be supported. This applies both to classroom courses as well as to self instruction (i.e. independent study) settings or hybrid forms thereof.

As a starting point for the planning of a new and/or adaptation of an existing course, different curricula are available in the Teradata University Network. Based on these materials, the lecture as well as the associated exercises, assignments, and even exams can be prepared. The numerous articles, book chapters, and research as well as practice reports available in TUN are particularly suitable. Furthermore, some PowerPoint slides as well as web seminars that can be used for self-instruction are provided in the Teradata University Network, allowing for the immediate adoption in the respective course. For preparing exercises, it is possible to revert to the case studies available in TUN. These can be worked on by the students individually or in groups. In most cases, the Teradata University Network provides the course instructors with teaching notes, questions, and sample solutions.

It is helpful to students to use the MicroStrategy BI Application Modules in connection with some exercises provided in TUN. In doing so, students gain hands-on knowledge and develop a thorough understanding of the lecture materials. In order to become familiar with the MicroStrategy tools, it is advisable for both instructors and students to complete the Micro Strategy tutorial and to do some lessons in the eTrainer. Within the tutorial, the user has the opportunity of giving reporting functions and navigation as well as the other features of the tool an unlimited and unrestricted trial (Wixom, 2004). The eTrainer is a web-based training environment that allows users to learn about handling the Micro Strategy tool and checking their acquired knowledge. Topics covered include basic tool navigation, looking for a report, changing a report (both in form and content), storing and deleting a report, as well as exporting and printing documents (MicroStrategy, n.d.).

To accomplish database exercises, the SQL (Structured Query Language) Assistant tool can be used. This enables users to work on the Teradata database with adequate data sets in order to interactively learn how to develop SQL queries. Besides accessing pre-defined data sets that are taken from popular database text books, it is also possible to upload individual data sets into the Teradata database. For the tailored use of the SQL Assistant tool, course instructors have the opportunity to create and administrate dedicated course environments. Password protection allows differentiation between databases used for in-lecture demonstrations and exercises, databases needed for assignments, and databases restricted to examinations. Consequently, Teradata University Network and Teradata Student Network users are granted both reading and writing access on the databases assigned to them.

It is worth noting that some of the aforementioned resources (e.g. PowerPoint presentations, case studies, and exercises) contain teaching notes which further simplify the development of a data warehousing or business intelligence course. If this assistance should prove insufficient and/or students have questions relating to Teradata University Network resources, course instructors may contact the Teradata customer service. Within the scope of TUN support service level agreements, inquiries are generally answered within a timeframe of 24 hours.

## **Application Case**

An application example for basing a distance learning course on Teradata University Network materials is the Management Information Systems (MIS) course that we teach every other semester as a core course in the Master of

Business Informatics program at the Virtual Global University (VGU, n.d.). We show which TUN resources were initially used in the conceptualization of the course, and which Teradata University Network/Teradata Student Network resources are available to students who attend the course. We also draw conclusions and present some lessons learned based on our experience as course instructors as well as on the students' feedback. Finally, we briefly report findings from two comparable courses on data warehousing and business intelligence that we teach once every academic year at the University of St. Gallen, Switzerland as an elective in the "Master of Arts in Information, Media and Technology Management" program (Fleisch & Gebauer, 2004).

## Course Outline and Objectives

The MIS course is intended to enable graduate students to understand the various types of information systems (IS) in commercial organizations as well as public administration. In contrast to the broad interpretation of MIS that is commonly followed at US universities (e.g. incorporating hardware and software foundations, MIS development, societal impacts, cultural issues and more), a narrow interpretation of MIS is followed in this course: MIS is defined as the various types of information systems used by commercial organizations and public administration in order to support management processes. As a consequence, the MIS course presents two different views for classifying MIS (business view, sessions 3 and 4; systems architecture view, sessions 5 and 6), as well as examples for every information systems sub-type. Additionally, analytical information systems are covered in more detail (sessions 7 and 8). As a conceptual foundation, information integration is presented as one of the primary goals of organizational information management. Students are provided with an overview of the changing scope and role of MIS from stand-alone systems to integrated components of information management and systems architecture (sessions 1 and 2). To round off and conclude the course, the most important management issues for information systems are presented, and future developments in the MIS field are outlined (sessions 9 and 10).

Since the sessions of the MIS course are conceptualized as self-instruction lessons, the main component of the course is a set of hypertext documents comprising the lecture notes and links to video presentations. These video presentations combine a video stream showing the instructor lecturing, with the associated PowerPoint slides of the presentation comprising the lecture material. The video presentations were produced by means of the PowerPoint add-on Microsoft Producer for Microsoft PowerPoint (Shelly et al., 2005). The lecture notes and video presentations are accompanied by links to assignments (either case studies or tool exercises), and recommended readings.

The topics and materials covered in the course and the associated assignments are exhibited in *Table 1*. The Teradata University Network resources that have been used for preparing the course are presented in the subsequent section.

*Table 1. Topics of lectures and assignments*

Session(s)	Topics covered in lecture(s)	Materials covered in assignment(s)
1	Introduction: definitions and foundations <ul style="list-style-type: none"> <li>– Data vs. information vs. knowledge</li> <li>– Information systems</li> <li>– Management information systems</li> <li>– Information management</li> <li>– Information integration</li> </ul>	Case study: "The benefits of data warehousing at Whirlpool" (Haley et al., 1999)
2	Changing scope of IS and different views on IS <ul style="list-style-type: none"> <li>– Information integration: directions, dimensions, techniques</li> <li>– Changing role and scope of IS</li> <li>– Different approaches to IS understanding <ul style="list-style-type: none"> <li>• Business view of IS</li> <li>• Systems architecture view of IS</li> <li>• Technical view of IS</li> <li>• Functional view of IS</li> </ul> </li> </ul>	Case study: "Data warehouse governance at Blue Cross and Blue Shield of North Carolina" (Watson et al., 2002)  Case study: "Implementing SAP R/3 at the University of Nebraska" (Sieber et al., 1996)
3 and 4	The business view on information systems <ul style="list-style-type: none"> <li>– Derivation of IS types from a business</li> </ul>	Case study: "Brittany Ferries makes plain sailing of using the Internet to deliver financial

	<p>viewpoint</p> <ul style="list-style-type: none"> <li>– Most important IS types <ul style="list-style-type: none"> <li>• Transaction processing systems</li> <li>• Office automation systems</li> <li>• Management information systems</li> </ul> </li> <li>– Other IS types <ul style="list-style-type: none"> <li>• Knowledge and information management systems</li> <li>• Decision support systems</li> <li>• Creativity support systems</li> <li>• Executive information systems</li> <li>• Inter-organizational systems</li> </ul> </li> </ul>	<p>statements to agents and resolve queries online” (AXS-one, 2005)</p> <p>Tool exercise: “AdVent Technology – Using the MicroStrategy Sales Analytic Module” (Bonney, 2003)</p> <p>Case study: “WISDOM provides competitive advantage at Owens &amp; Minor” (Stoller et al., n.d.)</p>
5 and 6	<p>The systems architecture view on information systems</p> <ul style="list-style-type: none"> <li>– Systems architecture model <ul style="list-style-type: none"> <li>• Model dimensions and resulting integration dimensions</li> <li>• Application types</li> <li>• Interface system types and reference model</li> </ul> </li> <li>– IS types and interface types from a systems architecture viewpoint <ul style="list-style-type: none"> <li>• Vertical applications</li> <li>• Information-centered applications</li> <li>• Analytical applications</li> <li>• Data warehouse systems</li> <li>• Horizontal applications</li> <li>• Enterprise application integration systems</li> <li>• Inter-company integration systems</li> </ul> </li> </ul>	<p>Case study: “Data warehousing supports corporate strategy at First American Corporation” (Cooper et al., 2000)</p> <p>Case study: “Continental Airlines Takes Off with Real-time Business Intelligence” (Anderson-Lehman et al., 2004)</p> <p>Case study: “Harrah’s high payoff from customer information” (TDWI, 2000)</p>
7 and 8	<p>Analytical applications</p> <ul style="list-style-type: none"> <li>– Reporting systems</li> <li>– Managed query environments</li> <li>– Decision support systems <ul style="list-style-type: none"> <li>• Decision support systems using artificial intelligence techniques</li> <li>• Data mining systems</li> <li>• Group decision support systems</li> </ul> </li> <li>– Executive information systems</li> <li>– On-line analytical processing and data marts</li> <li>– Support of process portals and packaged applications</li> </ul>	<p>Case study: “Production planning and control in textile industry” (Karacapilidis &amp; Pappis, 1996)</p> <p>Tool exercise: “On-line analytical processing”, using a PowerPlay web front end to explore data of the NASA workforce (NASA, n.d.)</p> <p>Case study: “CRM case study – Optimizing relationships at National Australia Bank, Inc.” (Khirallah, 2001)</p>
9	<p>Management of integrated information logistics</p> <ul style="list-style-type: none"> <li>– Three views on information management</li> <li>– Information systems management <ul style="list-style-type: none"> <li>• IS strategy and justification</li> <li>• IS architecture and portfolio planning</li> <li>• IS project management and support</li> </ul> </li> <li>– Data warehousing management <ul style="list-style-type: none"> <li>• Data warehousing strategy and justification</li> <li>• Data warehousing processes and tools</li> <li>• Data warehousing organization</li> </ul> </li> </ul>	<p>Case study: “3M moves to a customer focus using a global data warehouse” (Goodhue &amp; Wixom, 2001)</p>

10	<p>Trends in data warehousing and information logistics</p> <ul style="list-style-type: none"> <li>– What has been achieved so far...</li> <li>– Some current trends <ul style="list-style-type: none"> <li>• Inter-company information systems</li> <li>• Consideration of privacy legislation and concerns</li> <li>• Convergence of data warehousing and knowledge management</li> <li>• Using the web to create information</li> <li>• “Commodization” of data warehousing</li> </ul> </li> </ul>	Case study: “Analyzing the footsteps of your customers” (Theusinger & Huber, n.d.)
----	---	--

### Usage of Teradata University Network Resources

Various case studies and assignments have been adopted as they stand from Teradata University Network/Teradata Student Network (see *Table 1*). The pre-defined questions suggested by the case study and assignment authors were mainly adopted, and in some instances supported and/or substituted by individually designed questions. These referred to the case study itself and/or to the material covered in the respective lecture(s). The extensive teaching notes provided in TUN support the instructors in preparing elaborate reference solutions for the assignments. This information can be used both for grading as well as for giving the students well-founded feedback on their individual solutions.

In particular, the following case studies were taken from Teradata University Network/Teradata Student Network resources:

- “The benefits of data warehousing at Whirlpool” (Haley et al., 1999): The Whirlpool case illustrates a taxonomy of benefits that can be realized through data warehousing. Topics covered include: Data mart, data warehouse architecture, decision support systems and business intelligence, enterprise data warehouse, extraction-transformation-loading, on-line analytical processing, organizational issues, and return-on-investment.
- “Data warehouse governance at Blue Cross and Blue Shield of North Carolina” (Watson et al., 2002): This case study describes the award winning data warehouse governance practices at Blue Cross and Blue Shield of North Carolina. It also describes the warehouse, training and support, and the benefits realized. Topics covered include: Data governance, data models and modeling, data quality, data warehouse justification, extraction-transformation-loading, information requirements justification, and training and support.
- “Data warehousing supports corporate strategy at First American Corporation” (Cooper et al., 2000): This case describes the implementation of a customer relationship management strategy at First American Corporation. Topics covered include: customer relationship management, data mining, data quality, data warehouse justification, decision support systems/business intelligence, development methodologies, enterprise data warehouse, extraction-transformation-loading, metadata, on-line analytical processing, organizational issues, and project management.
- “Continental Airlines takes off with real-time business intelligence” (Anderson-Lehman et al., 2004): This case describes the development of real-time business intelligence at Continental Airlines. Powered by an active data warehouse, Continental has changed the way it does business. The case describes the conditions at Continental that led to this change, the real-time applications and technology that are used, the benefits realized, and the lessons that have been learned. Topics covered include: customer relationship management, data warehouse success/failure, decision support systems/business intelligence, enterprise data warehouse, extraction-transformation-loading, organizational issues, return-on-investment, and business impact.
- “Harrah’s high payoff from customer information” (TDWI, 2000): This case study describes Harrah’s customer relationship management business strategy, the technological and organizational changes that were required, the application of closed loop marketing, the resulting business benefits, and the lessons learned. Topics covered include: customer relationship management, data governance, data warehouse architecture, data warehouse success/failure, database performance, enterprise data warehouse, extraction-transformation-loading, operational data store, organizational issues, and return-on-investment.

- “3M moves to a customer focus using a global data warehouse” (Goodhue & Wixom, 2001): This case study describes the development of a data warehouse at 3M. It emphasizes the justification process, the technical infrastructure, and benefits. It is especially interesting because the savings from the consolidation of multiple decision support data marts exceeded the cost of the warehouse. Topics covered include: customer relationship management, data governance, data warehouse architecture, data warehouse justification, enterprise data warehouse, extraction-transformation-loading, hardware/software requirements, information requirements determination, organizational issues, and return-on-investment.

In addition to the aforementioned case studies, the Teradata University Network/Teradata Student Network tool exercise “AdVent Technology – Using the MicroStrategy Sales Analytic Module” (Bonney, 2003) is used in the MIS course. The exercise requires students to use the MicroStrategy 7 Sales Analytic Module (see Figure 2) to answer a total of three questions by submitting a screenshot or print-out of the produced reports. Students can access the MicroStrategy business intelligence front end by means of a standard web browser via Teradata Student Network. As already mentioned, it is advisable to require students to complete the MicroStrategy tutorial before working on the actual assignment.

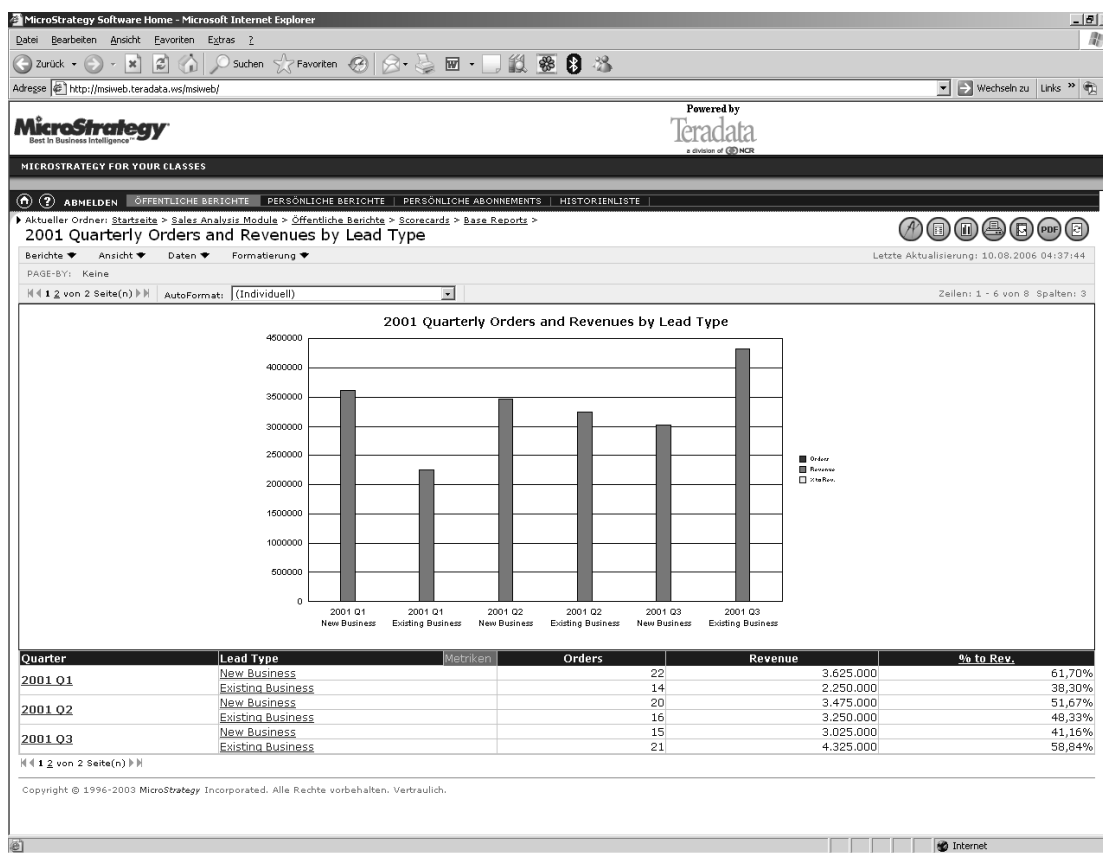


Figure 2. MicroStrategy Sales Analytic Module—Sample screenshot

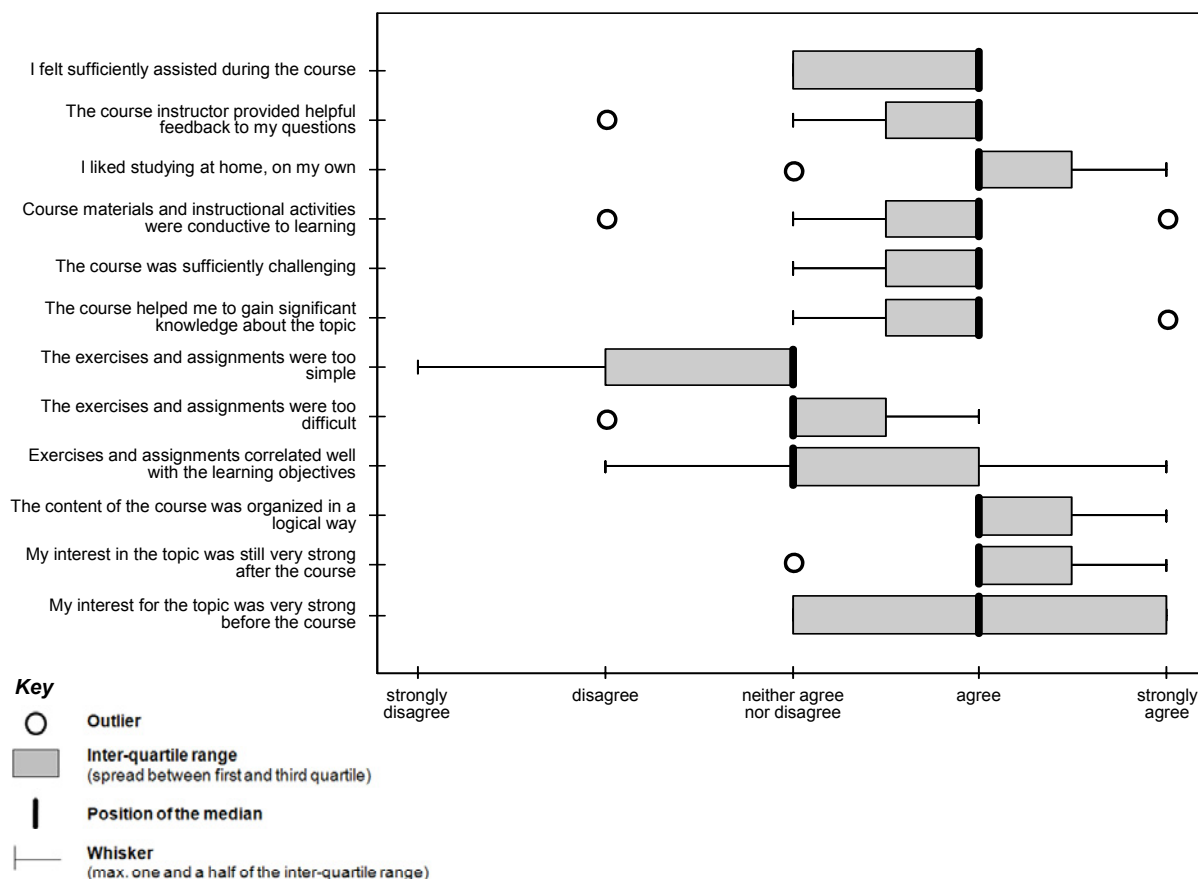
## Students' Feedback

At the end of the MIS course, students have the opportunity to provide the instructors with detailed feedback on the course contents and design by means of a standardized questionnaire. The questions cover statistical data, technical aspects (accessibility and quality of course materials, electronic communication with the instructors), evaluation of the course contents, and overall feedback and general comments.

Students' feedback on the MIS course has been very positive. All students who submitted the feedback questionnaire from the course taught in 2005/2006 characterized the overall course design as "good". Almost 60% of those reported "very good" or even "excellent" course design.

The box-whisker plot illustrated in *Figure 3* is a graphic representation of a set of descriptive statistics on the students' feedback on the 2005/2006 MIS course. The students were asked to indicate their level of agreement with twelve pre-specified statements (vertical axis of *Figure 3*) on a five-tiered Likert scale (horizontal axis of *Figure 3*). The underlying research question can be formulated as follows: "How do the students perceive the contents and design of the MIS course?"

Our initial hypotheses are that the self-instruction format as well as the electronic course materials, case studies, and tool exercises foster a positive learning environment and contribute to the overall learning success. The data from the students' feedback was analyzed with respect to these hypotheses. The length of the boxes in the box-whisker plot is determined by the respective inter-quartile range. The inter-quartile range is a measure for the statistical spread between the first and the third quartile, this covers 50% of the data. The position of the median within the inter-quartile range is depicted by the bold vertical lines within the boxes, indicating the skewness of the underlying distributions. The horizontal lines on the left and/or right sides of the boxes are referred to as whiskers. The length of the whiskers is restricted to one and a half of the respective inter-quartile range at the most, and in either case determined by an actual value of the data. Values outside of this range are separately marked by circles and referred to as outliers. If there are no values beyond the whiskers, their length is determined by the minimal and/or maximal value represented in the underlying data set.



*Figure 3.* Students' feedback on the MIS course



From the information depicted in the box-whisker plot, we conclude that the mainly electronic communication and interaction between students and instructors does not constitute any substantial barrier to learning. On the contrary, the majority of students felt sufficiently assisted and supported during the course, and liked the self-instruction format. It provided them with the opportunity to study on their own without any restrictions with respect to timing and/or study location. Students also perceived the electronic course materials and instructional activities as conducive to learning, and valued the logical organization of the contents. Regarding the level of difficulty, the majority of students characterized the course as “sufficiently challenging” and referred to the exercises and assignments as neither too simple nor too difficult. Our initial hypotheses about the positive effects of the self-instruction format as well as of the teaching materials therefore hold true.

However, the correlation between exercises and assignments on one side and learning objectives on the other side meets with slight criticism (cf. statement “exercises and assignments correlated well with the learning objectives”). Some students also commented on the partial lack of correspondence between assignments’ contents and the topics covered in the lectures. Clear needs for improvement exist: Either the lecture materials must be prepared in order to match the assignments, or the assignments have to be re-examined so they correspond with the lecture materials. However, since the Teradata University Network provides instructors with detailed teaching notes for the exercises and case studies, the instructors’ effort that is necessary for the adaptation and/or matching of these components is considered to be rather small.

Beside these points, overall feedback on the MIS course has been very encouraging. The majority of students agreed that the course helped them to gain significant knowledge about MIS, and reported that their interest in the topic remained equally strong or increased during the course.

### **Lessons Learned from Comparable Courses**

As already mentioned, we are also involved in the teaching of two comparable courses at the University of St. Gallen, namely one self-instruction course on business intelligence, and an in-classroom course on management information systems.

The business intelligence course covers the following topics: data warehousing architecture, data population (extraction-transformation-loading), data organization, database systems, Structured Query Language (SQL), multi-dimensional modeling, on-line analytical processing, data mining, performance management, customer relationship management, and other analytical applications. Similar to the MIS course previously described, the business intelligence course’s lecture material is made available to the students via a dedicated course web site. Students of the business intelligence course are also required to complete individual assignments and tool exercises that are based on and/or adopted from the Teradata University Network/Teradata Student Network.

In addition to the exercises outlined above, students have to complete exercises on multi-dimensional modeling and SQL. Teradata University Network/Teradata Student Network offers a wide range of corresponding assignments (e.g. Garfield, 2003 and 2004). The SQL queries on the Teradata database have to be performed by means of the Teradata SQL Assistant tool (see Figure 4). An introductory guide to SQL using this tool is provided in Teradata University Network/Teradata Student Network (Hoffer, 2005).

The students’ feedback on the assignments and especially on the tool exercises has been very positive. Students value the high quality of the case studies and see a huge advantage in gaining hands-on knowledge of the lecture materials by working with tools and tool demos provided in the Teradata Student Network.

The same is true for the course on management information systems held at the University of St. Gallen. This course covers basically the same material as the aforementioned MIS course, but is taught in-classroom. The students need to both attend the lectures and complete a total of eight exercises—seven tool exercises using (1) the Teradata SQL Assistant, (2) a multi-dimensional information modeling tool, (3) the MicroStrategy reporting modules, (4) an online business intelligence tool, (5) an open source data mining tool, (6) Microsoft Excel for decision support exercises, and (7) an executive information system adoption simulation tool as well as an integrative case study.

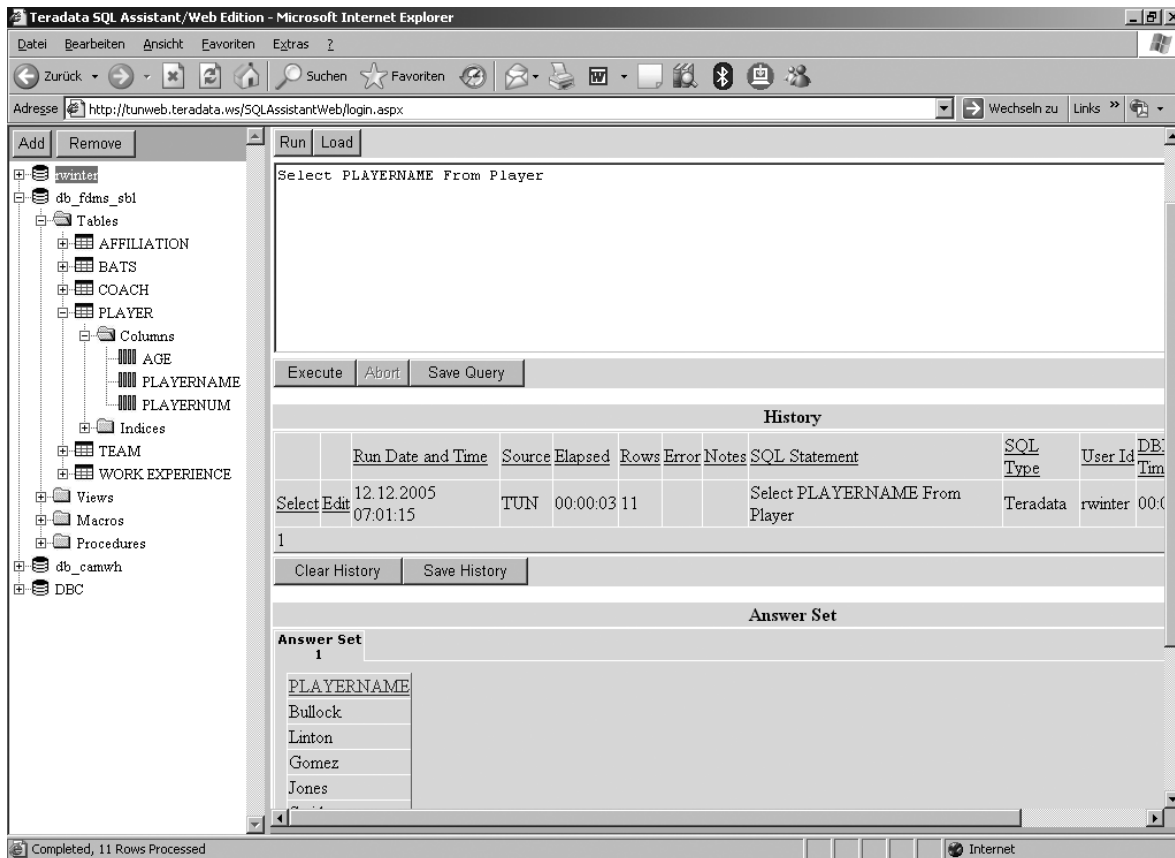


Figure 4. Teradata SQL Assistant tool—Sample screenshot

Teradata University Network/Teradata Student Network unfortunately does not contain a data extraction, transformation and loading tool nor a dedicated data mining tool. Since these two tool types are important for management information systems/business intelligence courses, an extension in this direction is advised.

## Limitations of the Teradata University Network

The advantages of the Teradata University Network can be seen in the easy exchange of teaching materials and in the free-of-charge access to state-of-the-art commercial software tools, but the limitations of this platform have to be taken into consideration as well.

Firstly, it has to be recognized that TUN will remain interesting for academia only if the available material is up-to-date. Secondly, it must be noted that the provided material is still fairly limited. These two facts indicate that it is necessary for lecturers and companies to remain willing to make their teaching materials and software tools available free of charge to others by using TUN.

The cost of the development and maintenance of the Teradata University Network and the Teradata Student Network are covered by software vendors. These companies aim at a marketing impact on future executives and lecturers. As a consequence, the choice of software tools provided on TUN is a result of Teradata's business and partnership strategy. So the amount of tools is limited to one tool for each field of application. However, a limited number of tools facilitates the integration of data sets and supports integrative application examples. In our courses, the SQL assignment, the information modeling assignment, and the reporting assignment operate on the same physical data set. Students therefore have the opportunity to learn about different analysis and visualization techniques without changes in meaning and context of the data and the depicted information.

Finally, a fourth limitation can be identified. The perspectives adopted within the Teradata University Network and the Teradata Student Network may be insufficiently diversified to suit many potential users. That is why validated findings about the actual use and the needs of the users are necessary; these would allow a focused evolution of the platform. Besides figures about registered Teradata University Network users and aggregated Teradata Student Network usage statistics, additional information regarding usage is currently collected in the Teradata Student Network. This data and the corresponding findings will be made available in the future.

## Conclusion and Outlook

The Teradata University Network is an innovative and effective support instrument for educators in the field of data warehousing, business intelligence, and database and the application of such information systems for management support. By supporting the development and execution of state-of-the-art, in-class as well as distance learning courses in these fields, the Teradata University Network helps to prepare students in a realistic, yet theory-grounded way. A substantial amount of teaching material (assignments, presentations, case studies, syllabi, etc.) is made available by an integrated platform with easy, free-of-charge access to state-of-the-art commercial software.

Useful extensions of TUN can be seen in the allocation of further teaching materials and software tools. The addition of a data mining tool, customer relationship management software, and an extraction-transformation-loading tool are planned. However, the integration of these tools as well as their evolution into a web-based application service provider model may take some time. After the planned additions, the Teradata University Network will provide a complete range of the most important types of data warehousing/business intelligence tools—but limited to one tool for each type.

Thanks to an increased internationalization of curricula and teaching aids, significant demand for non English-language materials has not been requested as yet. The multiple indexing of the content as well as the search functions make the Teradata University Network easier to reuse, entire courses as well as components. In order to support the integration of teaching materials into a broad range of courses and programs, the creation of even smaller content components and additional, searchable content metadata has been planned.

## References

- Anderson-Lehman, R., Watson, H. J., Wixom, B. H., & Hoffer, J. A. (2004). *Continental Airlines Takes Off with Real-time Business Intelligence*, retrieved June 19, 2008 from <http://kazman.shidler.hawaii.edu/ContinentalCase.pdf>.
- AXS-one (2005). *Case Study - Brittany Ferries. Brittany Ferries Makes Plain Sailing of Using the Internet to Deliver Financial Statements to Agents and Resolve Queries Online*, retrieved June 19, 2008 from <http://www.axsone.com/pdf/BrittneyFerries032602.pdf>.
- Baumgartner, P., Häfele, H., & Maier-Häfele, K. (2005). *Evaluation von Lernplattformen: Verfahren, Ergebnisse und Empfehlungen (Version 1.3)*, retrieved June 19, 2008 from [http://www.campussource.de/opensource/docs/evaluation\\_lernplattf\\_at.pdf](http://www.campussource.de/opensource/docs/evaluation_lernplattf_at.pdf).
- Bonney, L. (2003). *AdVent Technology: Using the MircoStrategy Sales Analytic Module*, retrieved June 19, 2008 from <http://www.teradata.com/t/page/135425/index.html>.
- Cooper, B. L., Watson, H. J., Wixom, B. H., & Goodhue, D. L. (2000). Data Warehousing Supports Corporate Strategy at First American Corporation. *MIS Quarterly*, 24 (4), 547-567.
- Fleisch, E., & Gebauer, H. (2004). *Das HSG Master-Programm für Business Innovation*, retrieved June 19, 2008 from [http://www.logistik.unisg.ch/org/logm/web.nsf/SysWebRessources/Information+IMT+Master+/\\$FILE/vorstellung\\_IMT\\_Master\\_Uni\\_St\\_Gallen.pdf](http://www.logistik.unisg.ch/org/logm/web.nsf/SysWebRessources/Information+IMT+Master+/$FILE/vorstellung_IMT_Master_Uni_St_Gallen.pdf).
- Garfield, M. (2003). *Six Data Modeling Exercises*, retrieved June 19, 2008 from <http://www.teradata.com/t/page/135748/index.html> (available after having logged in to TUN).

- Garfield, M. (2004). *Three Normalization Problems*, retrieved June 19, 2008 from <http://www.teradata.com/t/page/135762/index.html> (available after having logged in to TUN).
- Goodhue, D. L., & Wixom, B. H. (2001). *3M Moves to a Customer Focus Using a Global Data Warehouse*, retrieved June 19, 2008 from <http://www.teradata.com/t/page/135421/index.html> (available after having logged in to TUN).
- Graham, C. (2005). *Gartner Clients Consider Business Intelligence to Be a Major Priority*, retrieved June 19, 2008 from [http://www.gartner.com/DisplayDocument?ref=g\\_search&id=486797](http://www.gartner.com/DisplayDocument?ref=g_search&id=486797).
- Haley, B. J., Watson, H. J., & Goodhue, D. L. (1999). *The Benefits of Data Warehousing at Whirlpool*, retrieved June 19, 2008 from <http://www.igi-pub.com/downloads/teaching/IT5536.pdf>.
- Hoffer, J. A. (2005). *SQL Lab Manual*, retrieved June 19, 2008 from <http://www.teradata.com/t/page/135752/index.html> (available after having logged in to TUN).
- Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2004). *Modern Database Management* (7<sup>th</sup> Ed.), Upper Saddle River: Prentice Hall.
- Karacapilidis, N. I., & Pappis, C. P. (1996). Production Planning and Control in Textile Industry: A Case Study. *Computers in Industry*, 30 (2), 127-144.
- Khirallah, K. (2001). *CRM Case Study: Optimizing Relationships at National Australia Bank, Ltd.*, retrieved June 19, 2008 from [http://www.teradatalibrary.com/pdf/towergroup\\_020701b.pdf](http://www.teradatalibrary.com/pdf/towergroup_020701b.pdf).
- MicroStrategy (n. d.). *Overview of the MicroStrategy Site*, retrieved June 19, 2008 from [http://msi8web.teradata.ws/msidocs/MSTR8\\_TUN\\_Overview.pdf](http://msi8web.teradata.ws/msidocs/MSTR8_TUN_Overview.pdf).
- NASA (n. d.). *Exploring the NASA Workforce*, retrieved June 19, 2008 from <http://wicn.nssc.nasa.gov/index.htm>.
- Shelly, G. B., Cashman, T. J., & Sebok, S. L. (2005). *Microsoft Producer 2003: Essential Concepts and Techniques* (Vol. 1), Boston: Course Technology.
- Sieber, T., Siau, K., Nah, F., & Sieber, M. (1996). *Implementing SAP R/3 at the University of Nebraska*, retrieved June 19, 2008 from <http://homepages.strath.ac.uk/~cias19/teaching/40318/nebraskaU.pdf>.
- Stoller, D., Wixom, B. H., & Watson, H. J. (n. d.). *WISDOM Provides Competitive Advantage at Owens & Minor*, retrieved June 19, 2008 from <http://terry.uga.edu/~hwatson/owens&minor.doc>.
- TDWI (2000). *Harrah's High Payoff from Customer Information*, retrieved June 19, 2008 from <http://www.terry.uga.edu/~hwatson/Harrahs.doc>.
- Theusinger, C., & Huber, K.-P. (n. d.). *Analyzing the Footsteps of Your Customers*, retrieved June 19, 2008 from [http://ai.stanford.edu/~ronnyk/WEBKDD2000/WEBKDD2000\\_ARCHIVE/ii\\_2.pdf](http://ai.stanford.edu/~ronnyk/WEBKDD2000/WEBKDD2000_ARCHIVE/ii_2.pdf).
- VGU (n. d.). *Management Information Systems*, retrieved June 19, 2008 from <http://www.vg-u.de/cours1.asp>.
- Watson, H. J., Fuller, C., & Ariyachandra, T. (2002). *Data Warehouse Governance at Blue Cross and Blue Shield of North Carolina*, retrieved June 19, 2008 from <http://kazman.shidler.hawaii.edu/BCBSscase.pdf>.
- Watson, H. J., & Hoffer, J. A. (2003). Teradata University Network: A New Resource for Teaching Large Databases and Their Applications. *Communications of the AIS*, 12, 131-144.
- Watson, R. T. (2003). *Data Management: Databases and Organizations* (Vol. 4), New York: Wiley.
- Winter, R. (2004). Teradata University Network. In Leino, T., Saarinen, T., & Klein, S. (Eds.), *Proceedings of the 12<sup>th</sup> European Conference on Information Systems*, Turku, Finland, 1-4 (CD-ROM).
- Wixom, B. H. (2004). Business Intelligence Software for the Classroom - MicroStrategy Resources on the Teradata University Network. *Communications of the AIS*, 14, 234-246.
- Wixom, B. H., & Watson, H. J. (2001). An Empirical Investigation of the Factors Affecting Data Warehousing Success. *MIS Quarterly*, 25 (1), 17-41.

## Hypergraphics for history teaching - Barriers for causal reasoning about history accounts

**José M. Méndez and Manuel Montanero**

Facultad de Educación, Universidad de Extremadura, Campus universitario, 06071 Badajoz, Spain //  
mmontane@unex.es // Tel:+34-651188477

### ABSTRACT

This paper investigates the uses of various kinds of hypermedia format for history learning, which specifically emphasizes on the role of causal reasoning about history accounts. Three different groups in the last school year of Secondary Education (aged 16) studied the same materials about the Discovery of America in three different formats: (a) linear text in paper, (b) conventional hypertext with a content structure in network, and (c) hypergraphic with an explicitly causal structure and guiding questions inserted in the causal connections. The results in this last group were better in almost every causal reasoning task, but no statistically significant differences were obtained. The design of graphic information and specific interrogations for causal reasoning in a hypergraphic format could exert some positive effects. However, this type of resources is not enough to succeed in implying students with little previous knowledge in self-explanation and review processes of their causal models about history accounts.

### Keywords

Computer assisted learning, Cognition and technology, Hypergraphic, Causal reasoning, History education, Secondary education.

### Introduction

Research has shown that we can obtain benefits from working with multiple texts (Jacott, López & Carretero, 1998; Perfetti, Britt & Georgi, 1995; Voss, Carretero, Kennet & Silfies, 1994). Such concurrent use of diverse sources of information allows students to replicate the working methods of expert historians and social scientists as they move from data collection to ordering, categorizing and interpreting the information at hand (Carretero & Limón, 1993). What makes History memorable and coherent is its understanding through the elaboration of temporal and causal connections between the events (Perfetti, Britt & Georgi, 1995). However, the elaboration of causal explanations from different texts constitutes a very complex task demanding a strategic behaviour too exigent for most of the pupils (Rouet, Marron, Perfetti & Favart, 1998). Due to this reason, some studies have also explored the use of hypertext as a medium to make easier a more effective information management, linking multiple documents. Cognitive Flexibility Theory offers studied the kinds of navigational tools and linkages that can support effective student learning from hypermedia materials (Jacobson & Spiro, 1995; Jacobson, Maouri, Mishra & Kola, 1996; Spiro & Jehng, 1990; Spiro, Feltovich, Jacobson & Coulson, 1991). Nevertheless, several studies have not found clear differences between hypertext and linear texts, regarding comprehension measures. Apparently, learning materials, which are frequently offered in hypertextual format, do not seem to make users modify in a real way their linear reading strategies. On the contrary, a hypertext reading involves new threats for expositive content comprehension, mainly due to the important cognitive load it demands from the readers. They must remember the exact location of every new piece of information in the links and nodes jungle, take decisions about where they should go and remember previously visited pages or links, what usually causes a feeling of *disorientation* (León, 1998; Plowman *et al.*, 1999).

As either the content or task complexity increases the student's ability to navigate the hypertext declines. In the specific field of History teaching, diverse research pathways have found that hypermedia systems provide very limited benefits to users, conditioned, to a large extent, by students browsing conducts (Shapiro, 1998) or the assessment task type (Britt, Rouet & Perfetti, 1996). Conclusions in these works agree in the necessity of using a design model reflecting the organization top-level structure of the documents between themselves, as a favouring factor of hypertext "usability" and effectiveness.

Some works which have analysed those features that make more comprehensible causal texts in Social Sciences agree in highlighting three elements. Firstly, the events should be related between themselves following a temporal order by means of an explicit organization. Secondly, the materials should clarify the characters or human groups'

intentions involved in the historical events. Finally, information about the causal connection between two events, especially when these are distant in the text, should be provided (Linderholm *et al.*, 2000).

Therefore, in addition to user's hypertextual browsing experience, the aids insertion to organize and ensure the coherence in the causal explanations of historical events, could constitute an effective medium to prevent disorientation and improve the didactic utility of hypertextual information (Rouet, Levonen, Dillon & Spiro, 1996). In principle, there are two resources that could perform effectively these functions in hypertextual materials:

- the information arrangement by means of graphics reflecting the content causal structure
- interactive aids, in order to orientate the searching of information and causal reasoning processes.

## **Causal Organization**

Hypermedia technology allows to organize comfortably information into non-linear typology structures that, in contrast to traditional text, emphasize the most relevant relationships between each segment. The most widely studied typologies have been the network and hierarchical structures (De Vries & De Jong, 1999). Against what some technologists seem to assume, there are no clear evidences proving that these ways of arranging information are really advantageous for the learning of most of the students.

Network formats, commonly used, in which the user can navigate through different documents do not facilitate learning (Mohageg's, 1992; Simpson & McKnight, 1990). They can even produce worse results than traditional and linear versions (McDonald & Stevenson, 1996).

Unlike network typologies, the hierarchical ones are usually organized around graphics or tables to make easier users' orientation (Kommers & Lanzing, 1998). A hypergraphic is the graphics counterpart of hypertext: a linkage between related information by means of a graphic image. Instead of clicking on a word, users click on an icon to jump to the related information.

A meta-analysis carried out by Chen & Rada (1996) about 23 experimental studies with hypermedia materials came to the conclusion that the graphics presence helping to visualize the structure of information contributed significantly to achieve a greater utility of such systems. However, other studies have found that the advantages of hierarchical or mixed typologies are only really significant with students with a good reading capability but small previous knowledge of the content (Calisir & Zurel, 2003).

The possibilities for hypermedia materials are not reduced to the above mentioned typologies. In the case of History for example, the representation of a historical phenomenon is not necessarily hierarchical, neither must be distributed as an associative net. It should also reflect the causal connections between the distinct events and conditions of the historical moment studied. As we have already explained, the teaching material should provide an explicit arrangement of those events related between themselves, following a temporal and causal order.

## **Inferential implication resources**

A hypertext reading demands a great inferential effort to maintain information coherence (Stanton, Correia & Dias, 2000). In the case of History learning, the necessary inferences can be of a distinct type and complexity. Sometimes, we must wonder about the antecedents or causes of an event and sometimes about its consequences. Other inferences are mostly focused on the connections, that is, on understanding why an event is cause or consequence of another one. Finally, the difficulty of these causal inferences is presumably conditioned by the information degree of accessibility. Frequently, the readers are required to start from their prior knowledge to reconstruct that information not explicitly appearing in the sources or documents consulted. This last type of mental actions are the most important to achieve a deep comprehension of the model of the historical situation (Voss & Ney, 1996). However, limitations to understand the causal model do not reside exclusively in the lack of previous knowledge. Limitations at the moment of combining multiple causal relationships simultaneously in the work memory (Voss *et al.*, 1994) or the lack of active involvement on the part of the pupil in the causal explanation (Chi, 2000), constitute other obstacles which is necessary to bear in mind when creating hypermedia materials.

Previous research has looked into the benefits of embedding navigational questions and cues into hypertext, in order to support reasoning and metacognitive reflection on different subjects (Collis & Meeuwssen, 1999; Kashihar, Hiroshi & Toyoda, 1999; Relan & Smith, 1996; Sumner & Taylor, 1998; Veenman, Elshout & Busato, 1994). Apart from the organization and the hypergraphic aids, inserting *interrogations* making the pupil reflect about important aspects or concepts in the document could be an adequate strategy in this direction (Taylor, Alber & Walker, 2002). Such questions, not only should make the reader review and clarify the explicit information or recover relevant previous knowledge. More specifically, should constitute aids to guide the reasoning about those conditions and intentions, explaining causal connections between events. Otherwise, the pupil hardly could escape from the so-called “narrative trap” (Duthie, 1986), that is, from limiting to describe what happened, without really grasping why each of the events exposed were generated. The scheme provides the sequence of causes and consequences of the phenomenon but not the content justifying why those causes have their corresponding effects and not others. Then, the main point in comprehension lies in trying to infer explicitly *why* each event affects the others, in order to build and contrast a mental model of the situation. In this direction, the *guiding questions* inserted in precise places in the hypertext (such as the graphic connections between the events) could guide the reasoning and review process of the explicit and implicit information contained in the diverse documents. At the same time, they could favour a more active involvement on the part of the pupil.

The present study was directed to explore a hypertextual materials with these two resources: the hypergraphic representation of the multicausal structure and the insertion of guiding questions (which the user should answer throughout the process of study of the documents). On the one hand, we wanted to compare it with another type of hypertextual format in network, more conventional, and also in comparison with a linear presentation modality in paper. On the other hand, most of the research about hypertextual typologies have focused mainly on the speed and accuracy at the time of answering questions or locating explicit information in the documents. Our aim was to estimate the utility of these tools as a didactic resource for much more complex tasks, such as the multicausal historical reasoning.

## Method

### Participants

The sample consisted of 69 pupils (40 boys and 29 girls with an average age of 16.1), belonging to middle-lower classes, from a Secondary school. Their academic levels (i.e., vocabulary, knowledge, oral and written expression and comprehension levels) and motivation towards studies were equally low.

The subjects were randomly assigned to three conditions, related to the presentation format of a History content (about the Discovery of America): *Hypergraphic format with a multicausal organization* (Hypergraphic condition), *Hypertextual format with network organization* (Hypertext condition), and *Linear textual format* (Text condition).

The subjects’ *academic performance* in the linguistic and social sciences areas, the *previous knowledge* about the thematic content, and also the *degree of familiarity* with computer use and with navigation throughout hypertexts were taken into account to make sure of the initial homogeneity of the experimental groups.

### Intervention materials

a) *Hypergraphic format, with a multicausal organization* (acceding to the information by means of an arrow diagram) and *causal questions inserted in the graphic arrows* (Hypergraphic condition). The resultant expositive document was a hypertext whose main page reproduced a graphic showing the causal relationships between the different sections in which the theme was divided. A total of 19 expositive documents plus one used as a glossary of terms were elaborated. Each document had an average extension of 180 words (with documents ranging from 53 words to a maximum extension of 316 words) and 16 lines.

Navigation was performed through this graphic. Clicking on any of the graphic boxes, the document the titles of that event or condition referred to, appeared. Clicking on any of the arrows connecting boxes, questions making reference to the explicit and implicit relationships between those events, appeared too. These questions were answered in paper

by means of a written questionnaire consisting of 28 items with 4 possible answers. Three types of questions should be answered using explicit information (about the causes, consequences or a causal relationship nexus) and four using relatively implicit information (about the causes, consequences, causal relationship connection or creative reasoning).

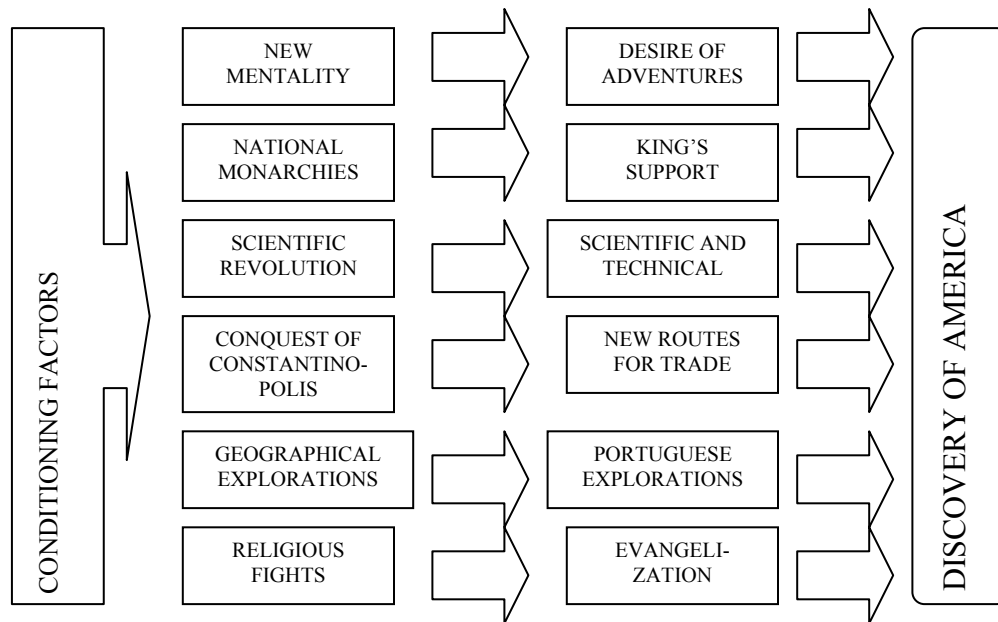


Figure 1. Interactive graphic aid for Hypergraphic condition

b) Conventional hypertextual format, with a non-interactive specific graphic and a brief questionnaire consisting of unspecific questions (Hypertext condition). It consisted of a conventional hypertextual document with a frame on the left where the epigraphs in which the document was divided appeared. Clicking on any of those links you can navigate to the headline of the epigraph at issue. Another way of navigating through the document was by means of the slider placed on the right of the screen, because the document could be read in a linear way downwards. At the end of the document the same causal scheme as in the previous hypermedia format was reproduced, but without the possibility of clicking on it to link with those sections it referred to.

Besides, in this format, the subjects answered a written questionnaire in paper consisting of unspecific questions not inserted in the content causal structure. In all, they had to answer 10 easy questions about dates, data or explicit ideas dealt in the reference document, similar to those we can usually find in text books at the end of each unit.

c) Linear textual format, with the same graphic and the same questionnaire used in the previous modality (Text condition). Unlike Hypergraphic and Hypertext conditions, the subjects did not have to read the information on a screen, but on the paper itself: a thirteen pages booklet where the same content and in the same order as those appearing in the conventional hypertextual format were reproduced. In the last page, the same causal scheme contained in the other formats appeared.

### Assessment materials

As *dependent variables*, different kinds of causal reasoning, exposed in Table 1, were considered. The assessment was based on a series of tasks which were applied immediately after the material study, except the delayed assessment test, made a week later.



Table 1. Causal Reasoning Assessment

Tasks	Causal inferences
Summary (immediate)	<ul style="list-style-type: none"> <li>• Quantity of causal factors remembered</li> <li>• Quantity of causal relationships remembered and explained</li> <li>• Global quality of the causal explanation</li> </ul>
Multicausal reasoning test (immediate)	<ul style="list-style-type: none"> <li>• Importance hierarchy of the diverse causal factors</li> <li>• Appraisal of the number of causal factors explaining a consequence</li> <li>• Elaboration of inferences about the causal nexus between two events (from explicit and implicit information)</li> </ul>
Explicit and implicit information inference test (delayed)	<ul style="list-style-type: none"> <li>• Elaboration of inferences about the causal factor or the consequence of an event (from explicit and implicit information)</li> <li>• Elaboration of inferences about the causal nexus between two events (from explicit and implicit information)</li> <li>• Elaboration of creative inferences about the situation model</li> </ul>

a) *Summary*. In the first task they should elaborate a summary as an explanation of the most important aspects studied and the causal relationship between them. Appraisal was based in two criteria. On the one hand, the number of factors and causal relationships remembered by each subject. On the other hand, the global quality of the causal explanation built by each pupil was qualitatively appraised. To do that, the historical causality appraisal scale made by Domínguez and Pozo (1998) was used. Basing on previous studies, these authors point out five levels in the appraisal of the quality of the causal explanations provided by the subjects when dealing with tasks of this nature:

- Level 1: absolute absence of the notion of cause as a factor relating facts between themselves. Facts or events appear as an unconnected set.
- Level 2: facts are inserted in an undifferentiated chain through a temporal sequence of such facts.
- Level 3: subjects seem to understand a certain notion of causality and they are able to recognize that a cause can affect more than one event or that a certain event can be affected by several causes at the same time. However, these relationships are not systematically explored or analysed.
- Level 4: subjects establish complete causal connections between the events, appreciating the interactions between them and its reciprocal relationships.
- Level 5: subjects, not also are able to establish a complete and reciprocal causal connection network between the events, but also to locate these in their corresponding spatial-temporal context.

b) *Multicausal reasoning test*. The student's multicausal reasoning capability was appraised by means of three convergent tasks:

- First of all, subjects were asked to establish a hierarchy with the degree of importance of each factor at the time of explaining the *Discovery of America*, reasoning their answers. Previously, three History teachers had estimated which should be the ideal hierarchy for this test.
- Secondly, they were asked to explain how many causal factors had intervened in their opinion and to explain the Discovery by means of these factors.
- Finally, to appraise the elaboration of implicit and explicit inferences, pupils answered a test whose questions made reference to the causal relationships and connections between certain contextual aspects of the epoch and the causal factors exposed, and also between the factors themselves.

c) *Delayed assessment test: implicit and explicit inferences*. The last assessment test was a questionnaire consisting of 24 items, each with four possible answers, demanding the elaboration of inferences about the causal factor, consequence or about the causal nexus between two events (taking as starting point the explicit and implicit information).

## Procedure

Subjects were randomly assigned to the three experimental conditions. Due to internal policies in the educative centre where the tests were carried out and due to the availability of use in the computers room, tests were performed

at different morning hours along different days throughout two weeks. The groups associated to the different experimental conditions carried out the tests simultaneously and in the same room, under the researcher's supervision. Two consecutive sessions were needed for the experimental session: 60 minutes were devoted for the consultation of the document about the Discovery of America (at the same time the pupils answered the questions in a self-evaluation questionnaire), and 30 minutes for the fulfilment of the immediate assessment tasks, without the presence of the document. A week later, during 30 minutes in a normal Social Sciences session, pupils answered to the questions contained in the delayed assessment test. The Social Sciences teachers provided no information about *the Discovery of America* to any of the groups involved neither in the previous weeks nor during the performance of the tests.

The degree of involvement in the task at the time of answering the different questionnaires, and also when consulting the documents of each experimental condition, was quite low. Many of the subjects did not answer those questions demanding a greater effort (open-ended questions) or those questions at the second half of the questionnaire just by demotivation. No subject showed difficulties when using hypergraphic or hypertextual formats, nor orientation problems at the time of consulting the different files of which the document was composed of. In general, subjects under Text condition finished earlier of consulting the whole document.

## Results

### Summary

A 37.5% of the subjects included in Hypergraphic condition, a 59.09% in Hypertext and a 47.82% in the case of Text condition, did not answer this test, that is, a 48% of all the subjects, what obviously reduces the interpretation of the results shown in Table 2. Most of the pupils left this test blank, not because they were not able to write, with higher or lower quality, a summary of what they had read, but, simply for unwillingness.

As to the quality of the summary, the level of the causal or historical explanations for most of the subjects would be between level 2 and level 3 of the scale described by Domínguez and Pozo (1998). Scarcely a third of the subjects was able to establish such causal relationship between some of the factors. The subjects included in Hypergraphic condition obtained higher average marks. At first sight, it seems that the fact of having the scheme with the causal structure permanently present, made easier for this modality to remember in a more efficient way the main contextual factors and elements of the theme. However, there were no significant differences between the three experimental conditions as to the quantity of causal factors and contextual elements remembered.

Table 2. Means and standard deviations in the summary test in the Hypergraphic, Hypertext and Text condition

<i>Summary</i>	<i>Modalities</i>	<i>M</i>	<i>SD</i>
Number of summarized factors	Hypergraphic	0.67	1.20
	Hypertextual	1.41	1.50
	Text	1.09	1.68
Number of summarized relationships	Hypergraphic	0.17	0.48
	Hypertextual	0.59	0.91
	Text	0.01	0.29
Global quality of the summary	Hypergraphic	2.12	2.15
	Hypertextual	0.59	1.33
	Text	1.17	1.75

### Multicausal reasoning test

As to the three multicausal reasoning tasks, we could observe again higher average marks in hypertext with interactive graphic aids condition (Table 3). Such difference in favour of those individuals grouped in Hypergraphic condition is shown in 9 out of the 13 questions which composed the two blocks of questions related to the elaboration of explicit and implicit causal inferences, though they were not statistically significant.

As to the question of how many and which factors would explain the Discovery, there were not differences between the subjects included in the three experimental conditions either. A third of the subjects did not answer the question. A 26% considered that several factors exerted influence (in general, those factors most of the subjects valued). A 20% considered that there was only one causal factor explaining the historical event. Finally, just a 20% considered that all the factors had an influence in a higher or lower degree.

This absence of significant differences repeated again in the question related to the appraisal of the importance of each causal factor. It is necessary to stand out the tendency to appraise those factors of an intentional or motivational kind (personal factors) over those of a structural kind.

*Table 3.* Means and standard deviations in the causal reasoning test and in the causal inference test in the Hypergraphic, Hypertext and Text condition

<i>Reasoning</i>	<i>Modalities</i>	<i>M</i>	<i>SD</i>
Hierarchy of factors (immediate)	Hypergraphic	1.75	1.85
	Hypertextual	0.91	1.15
	Text	1.22	1.56
Number of factors (immediate)	Hypergraphic	1.08	1.21
	Hypertextual	1.04	1.40
	Text	0.87	1.29
Causal inferences (immediate)	Hypergraphic	3.19	3.03
	Hypertextual	2.00	3.06
	Text	1.54	2.65
Information inference test (delayed)	Hypergraphic	6.30	4.33
	Hypertextual	4.18	3.84
	Text	4.96	5.02

### Information inference test

Subjects in Hypergraphic group got better marks than the rest in all the questions asking explicit and implicit information inference (table 3). The other two groups got very similar marks. Even though this could indicate that the graphic aid provided to the Hypergraphic group has helped to better identify the relationships between each causal factor and their contextual conditions, differences were not significant either.

If we analyse the success percentages in the different types of questions designed for this test, subjects included in Hypergraphic condition got better percentages than those included in Text group in all the questions types, and better than those included in Hypertext group in all the questions, except in those related to the causal nexus (Table 4).

*Table 4.* Success percentages in each of the inference types (delayed) in the Hypergraphic, Hypertext and Text condition

<i>Modality</i>	<i>Causal</i>	<i>Consecutive</i>	<i>Nexus</i>	<i>Creative</i>	<i>Explicit</i>	<i>Implicit</i>
Hypergraphic	43.7%	41.6%	34.35%	39.2%	39.08%	40.32%
Hypertextual	28.6%	34.8%	37.45%	35.6%	32.6%	34.8%
Text	32.12%	36.15%	26%	38.4%	32.2%	32.5%

### Discussion and conclusion

The main goal of our research was to compare the effect of different modalities for improving multicausal reasoning, combining different formats and aids. The results are congruent with those provided by the researchers mentioned in the introduction regarding the difficulty of showing significant differences as to the *hypermedia format* presentation with materials relatively complex (McKnight et al., 1992; Britt et al., 1996; Goldman, 1996; Shapiro, 1998; Baraab, Young & Wang, 1999; Potelle & Rouet, 2003).

However, it can be observed a trend toward obtaining better results in those pupils who studied the hypergraphic format, with a *multicausal organization* and with *guiding questions*, designed to favour causal inferences. This is clearly shown in the fact that those subjects included in Hypergraphic group got better average marks in the different causal reasoning tasks. We have also clues proving that the graphic aid (a multicausal diagram, used as a navigation map) made easier the reference actions of the information. So, in spite of the lack of experience on the side of the subjects as to what supposes to consult a complex information in a hypertext, these never showed orientation problems.

There are two main reasons to explain why we can not observe a clear influence of the materials used in the study. On the one hand, the hypergraphic resources designed could not be really efficient to promote causal reasoning. On the other hand, these resources could not be adapted to the navigation strategies and to the characteristics of the subjects who took part in the study.

As to the first question, it is possible that either the representation of the causal content or the design of the guiding questions did not constitute adequate resources to success in involving students in the causal reasoning process. Instead of providing an *organization and a graphic representation* already made, Masterman and Sharples (2002) found positive results with a software which allowed students to build their own causal graphic from the linear reading of a historical phenomenon. Besides categorizing the properties of each event and connect them causally, subjects had to explain and justify the graphic made behind their classmates, what made easier the reflection and review of their implicit conceptions. In the same direction, Ravenscroft (2000) has found positive effects in the causal explanations about a Physics content from subjects exposed to simulated situations of dialog by means of a computer's tool. On the contrary, in our case, pupils did not have the opportunity to re-build the representation of their own causal model, nor took part in discussion activities about such model, what probably could have improved the results in the causal reasoning tasks.

The *guiding questions*, could help to generate causal inferences about the causal connections. In fact, if we base on the data collected through the delayed assessment test, subjects in Hypergraphic condition get higher success percentages in almost all the types of questions. Sinatra, Beck y McKeown (1993) found that a group of students who, during the reading of several Social Sciences texts, had to answer to causal questions, got better results than those who read the original text and even those who read a reviewed version, but differences were not statistically significant either. A reason why the use of this type of mediators tends to produce only small gains could have to do with an argument affirming that "answering questions is useful to achieve the goals of others, instead of being useful to review the own comprehension problems" (Chi, 2000, p. 225). Those inferences generated by means of this aid in the Hypergraphic modality would only serve to complete pieces of information or to join unconnected pieces of it. On the contrary, readers would use the self-explanations to review and repair their mental model when they find a conflict between this and the information contained in the text. In this direction, the experiments designed by Chi (1994) encouraged pupils to formulate self-explanations while reading aloud. Those who elaborated a greater number of self-explanations, got higher levels of learning than those formulating an inferior number, and these, in their turn, got better results than those who simply read the text several times. A more powerful resource than that incorporated to the *hypermedia format* would be, therefore, to design inferential engagement activities in a *e-learning* environment. The previous results reveal the risk of granting such resources a self-sufficient character, within the technologic tool being implemented. More elaboration regarding inference, reasoning and graphics design are encouraged. In future research, it is necessary to explore its complementation with causal discussion activities, such as those which some History teachers usually perform.

On the other hand, the observation of the *reading strategies* followed by the subjects during the experiment, bring us another possible explanation about the results found. In a first reading of the documents, both the subjects in Hypergraphic condition and those in Hypertext followed a linear reading strategy, similar to that used by subjects in Text condition. Only after this first reading, subjects adopted a different strategy, mainly because they had to carry out a specific task (answering to the guiding test), skipping from some documents to others to find the information needed. The reading strategy can be, in consequence, hiding away the effect of the other variables over the results obtained by the subjects, as also has been observed in other studies (see Foltz, 1996, or Potelle & Rouet, 2003). A fact supporting this last analysis can be deduced from the tendency of the subjects in Hypergraphic condition to bring out "the desire for adventures" factor in a higher degree than the subjects belonging to the other two experimental conditions. This result can only be explained as a consequence of the linear reading strategy followed behind the navigation map found in the hypertext home page. The "desire for adventures" factor was located at the top of the

multicausal diagram (see again figure 1) and was the first of the causal factors subjects of this condition would read following a linear scheme (from top to bottom and from left to right).

Finally, there can be no doubt that navigating throughout the information contained in a hypertext involves a cognitive over-effort for the subject. The smaller is his/her reading level and experience in the use of hypertexts, the greater is the effort (McKight, Dillon & Richardson, 1990; Rouet et al., 1996). In this direction, another possible fact which would explain why *hypermedia formats* have had so limited effects could be related to the low academic performance of the sample. In fact, clear differences between the questions about explicit and implicit information were not found. The model which most of their summaries reflect would be “subject plus details”, that is, a group of unconnected ideas about diverse contextual factors or elements of the Discovery of America not showing the comprehension of the historical event such texts refer to. A low engagement in the task could influence on the results as many pupils did not answer those open questions due to unwillingness or tiredness. Having into account the importance of the personal involvement of the pupil in this type of tasks, it is crucial to do research work on these personal factors in later studies.

## Acknowledgements

This research has been financed by the Junta de Extremadura-Consejería de Educación de Ciencia y Tecnología and FEDER through the II Plan Regional de Investigación, Desarrollo Tecnológico e Innovación (reference number 2PR04A096).

## References

- Baraab, S. A., Young, M. F. & Wang, J. (1999). The effects of navigational and generative activities in hypertext learning on problem solving and comprehension. *International Journal of Instructional Media*, 26, 283-301.
- Britt, M.A., Rouet, J.F. & Perfetti, Ch. (1996). Using hypertext to study and reason about historical evidence. In J. F. Rouet, J. Levonen, A. Dillon & R. Spiro (Eds.) *Hypertext and Cognition*, Mahwah, N.J.: LEA, 43-72.
- Calisir, F. & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior*, 19, 135-145.
- Carretero, M. & Limón, M. (1993). Aportaciones de la psicología cognitiva y de la instrucción a la enseñanza de la Historia y las Ciencias Sociales. *Infancia & Aprendizaje*, 62-63, 153-167.
- Chen, C. & Rada, R. (1996). Interacting with Hypertext: A Meta-analysis of Experimental Studies. *Human Computer Interaction*, 11, 125-156.
- Chi, M. T. H. (2000). Self-explaining texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.) *Advances in Instructional Psychology*, Mahwah, N.J.: LEA, 161-238.
- Chi, M. T. H., De Leeuw, N., Chiu, M.H. & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.
- Collis, B. & Meeuwssen, E. (1999). Learning to learn in a www-based environment. In C. Johnson, C. Hale, D. French & G. Farr (Eds.) *Internet-based learning: an introduction and framework for higher education and business*, Sterling, Virginia: Stylus Publishing, 25-46.
- De Vries, E. & De Jong, T. (1999). The design and evaluation of Hypertext structures for supporting design problem solving. *Instructional Science*, 27, 285-302.
- Domínguez, J. & Pozo, J. I. (1998). Promoting the Learning of Causal Explanations in History through Different Teaching Strategies. In J.F. Voss & M. Carretero (Eds.) *Learning and Reasoning in History. International Review of History Education*, London: Woburn Press, 344-359.
- Duthie, J. (1986). The Web: a powerful tool for teaching and evaluation of the expository essay. *The History Social Science Teacher*, 21, 232-236.

- Foltz, P.W. (1996). Comprehension, Coherence and Strategies in Hypertext and Linear Text. In J.F. Rouet, J. Levonen, A. Dillon & R. Spiro (Eds.) *Hypertext and Cognition*, Mahwah, N.J.: LEA, 109-136.
- Jacobson, M. J., Maouri, C., Mishra, P. & Kolar, C. (1996). Learning with hypertext learning environments: theory, design, and research. *Journal of Educational Multimedia and Hypermedia*, 5 (3), 239-281.
- Jacobson, M.J. & Spiro, R.J. (1995). Hypertext learning environments, cognitive flexibility and the transfer of complex knowledge: an empirical investigation. *Journal of Educational Computing Research*, 12 (5), 301-333.
- Jacott, L, López-Manjón, A. & Carretero, M (1998). Generating Explanations in History. In J.F. Voss & M. Carretero (Eds.) *Learning and Reasoning in History*, London: Woburn Press, 294-306.
- Kashihara, A., Hiroshi, U. & Toyoda, J. (1999). Reflection support for learning in hyperspace. *Educational Technology*, 39 (5), 19-22.
- Kommers, P. & Lanzing, J. (1998). Mapas conceptuales para el diseño de sistemas de hipermedia. Navegación por la web y autoevaluación. In J. A. León & C. Vizcarro (Eds.) *Nuevas Tecnologías para el Aprendizaje*, Madrid: Pirámide, 14-27.
- León, J. A. (1998). La adquisición de conocimientos a través del material escrito: texto tradicional y sistema de hipertexto. In J. A. León & C. Vizcarro (Eds.) *Nuevas Tecnologías para el aprendizaje*, Madrid: Pirámide, 68-93.
- Linderholm, T., Everson, M. G., van den Broek, P., Mischinski, M., Crittenden, A. & Samuels, J. (2000). Effects of Causal Text Revision on More- and Less-Skilled Readers' Comprehension of Easy and Difficult Texts. *Cognition and Instruction*, 18 (4), 525-556.
- Masterman, L. & Sharples, M. (2002). A theory-informed framework for designing software to support reasoning about causation in history. *Computers and Education*, 38, 165-185.
- McDonald, S. & Stevenson, R. J. (1998). Effect of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, 40 (1), 18-27.
- McKnight, C., Dillon, A. & Richardson, J. (1990). A comparison of linear and hypertext formats in information retrieval. In R. McAleese & C. Green (Eds.) *Hypertext: State of the Art*, Oxford: Intellect Book Ltd, 10-19.
- Mohageg, M. F. (1992). The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors*, 34, 351-367.
- Perfetti, C., Britt, M.A. & Georgi, M. C. (1995). *Text-based Learning and Reasoning*, LEA: Mahwah, N.J.
- Perfetti, Ch., Britt, M.A., Rouet, J.F., Georgi, M. & Mason, R. (1994). How students use texts to learn and reason about historical uncertainty. In M. Carretero & J. F. Voss (Eds.) *Cognitive and Instructional Processes in History and the Social Sciences*, Mahwah, N.J.: LEA, 257-282.
- Plowman, L., Luckin, R., Laurillard, D., Stratfold, M. & Taylor, J. (1999). Designing Multimedia for Learning: Narrative Guidance and Narrative Construction, In the proceedings of *CHI 99*, Pittsburgh, PA USA: ACM, 310-317.
- Potelle, H. & Rouet, J.F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *Human-Computer Studies*, 58, 327-345.
- Ravenscroft, A. (2000). Designing argumentation for conceptual development. *Computers and Education*, 34, 241-255.
- Relan, A. & Smith, W.C. (1996). Learning from hypermedia: a study of situated versus endemic learning strategies. *Journal of Educational Multimedia and Hypermedia*, 5 (1), 3-21.
- Rouet, J.F., Levonen, J., Dillon, A. & Spiro, R. (1996). *Hypertext and Cognition*. LEA: Mahwah, N.J.
- Rouet, J.F., Marron, M.A., Perfetti, C. & Favart, M. (1998). Understanding Historical Controversies: Students' Evaluation and Use of Documentary Evidence. In M. Carretero & J. F. Voss (Eds.) *Learning and Reasoning in History*, London: Woburn Press, 95-116.
- Shapiro, A. M., (1998). Promoting Active Learning: The Role of System Structure in Learning from Hypertext. *Human Computer Interaction*, 13, 1-35.

- Simpson, A. & McKnight, C. (1990). Navigation in Hypertext: Structural cues and mental maps. In R. McAleese & C. Green (Eds.) *Hypertext: State of the Art*, Oxford: Intellect Book Ltd, 74-83.
- Sinatra, G. M., Beck, I. L. & Mckeown, M. G. (1993). How knowledge influenced two intervention designed to improve comprehension. *Reading Psychology: An International Quarterly*, 14, 141-163.
- Spiro, R.J., Feltovich, P.J., Jacobson, M.J. & Coulson, R.L. (1991). Cognitive flexibility, constructivism and hypertext: random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Psychology*, 31 (5), 24-33.
- Spiro, R.J. & Jehng, J. (1990). Cognitive flexibility and hypertext: theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R.J. Spiro (Eds.) *Cognition, education and multimedia: explorations in high technology*, Hillsdale, New Jersey: Lawrence Erlbaum, 163-205.
- Stanton, N, Correia, A.P. & Dias, P. (2000). Efficacy of a map on search, orientation and access behaviour in a hypermedia system. *Computers and Education*, 35, 263-279.
- Sumner, T. & Taylor, J. (1998). Media integration through media-learning environments. In M. Eisenstadt & T. Vincent (Eds.) *The knowledge web: learning and collaborating on the net*, London: Kogan Page, 63-78.
- Taylor, L.K., Alber, S. & Walker, D.W. (2002). The Comparative Effects of a Modified Self-Questioning Strategy and Story Mapping on the Reading Comprehension of Elementary Students with Learning Disabilities. *Journal of Behavioral Education*, 2 (2), 69-87.
- Veenman, M.V., Elshout, J. J. & Busato, V.V. (1994). Metacognitive mediation in learning with computer-based simulations. *Computers in Human Behaviour*, 10 (1), 93-106.
- Voss, J.F., Carretero, M., Kennet, J. & Silfies, L.N. (1994). The collapse of the Soviet Union: a case study in causal reasoning. In M. Carretero & J. F. Voss (Eds.) *Cognitive and Instructional Processes in History and the Social Sciences*, Mahwah, N.J.: LEA, 403-429.
- Voss, J. F. & Ney, L. (1996). Learning From History Text: The Interaction of Knowledge and Comprehension Skill With Text Structure. *Cognition and Instruction*, 14 (1), 45-68.
- Wright, G.H. (1991). *Explicación y Comprensión*, Madrid: Alianza.

## Developing an Intelligent Diagnosis and Assessment E-learning Tool for Introductory Programming

**Chenn-Jung Huang**

Institute of Learning Technology, National Hualien University of Education, Hualien, Taiwan  
cjhuang@mail.nhlue.edu.tw

**Chun-Hua Chen, Yun-Cheng Luo, Hong-Xin Chen**

Institute of Learning Technology, National Hualien University of Education, Hualien, Taiwan

**Yi-Ta Chuang**

Institute of CSIE, National Chiao Tung University, Hsinchu, Taiwan

### ABSTRACT

Recently, a lot of open source e-learning platforms have been offered for free in the Internet. We thus incorporate the intelligent diagnosis and assessment tool into an open software e-learning platform developed for programming language courses, wherein the proposed learning diagnosis assessment tools based on text mining and machine learning techniques are employed to alleviate the loading of the teachers. Experiments were conducted in two introductory-undergraduate programming courses to examine the effectiveness of the proposed diagnosis and assessment tools. The learners' work including the source code and comments were processed by the proposed text mining and machine learning techniques. This system also provides immediate feedback and high-quality evaluation results to guide the learners with poor performance. Our experimental results reveal that the proposed work can effectively assist the low-ability learners.

### Keywords

Text mining, Multimembership Bayesian classifier, Support vector machines, Diagnosis, Assessment, E-learning platform

### Introduction

In recent years, several e-learning platforms have been developed to aid students in learning programming language (Fix & Wiedenbeck, 1996; Takemura *et al*, 1999; Mungunsukh & Cheng, 2002; Hulls *et al.*, 2005). Most of them used hypermedia since its hypertext structure reflects a model of learning based on the students' semantic memory model. Fix & Wiedenbeck (1996) designed an intelligent tool to aid the students, who already have knowledge of another programming language, in learning Ada programming language. This tool provided associated teaching material to help the learners comprehend the new programming language. Although the teaching material was generated dynamically and the students can navigate different learning topic through hyperlink, the students still have difficulty in planning a solution. Although Mungunsukh & Cheng (2002) proposed a case based reasoning approach to diagnose students' programming skill by extracting the events caught by student's learning activity and giving useful explanation and suggestion, the effectiveness of case based reasoning approach is unverified. The comprehension states of the learners were measured by tests during three learning phases in a Java programming language e-learning platform (Takemura *et al*, 1999). However, there is little help for the ones that have obstacle in planning of the solution for the given programming exercises. Greyling *et al* (2006) proposed a programming support tool for introductory programming courses that tends to concentrate on the syntax of a programming language during program developing. Nevertheless, their system was lack of providing immediate feedback concerning the correctness of the designed program.

It is observed that developing the programs during coding phase is difficult for most programming novices. We thus employ text mining and machine learning techniques to develop a programming diagnosis and assessment tool for an e-learning platform in this work to give the learners the guidance based on student's learning portfolios, whenever the learner is confused or stalled in programming. The guidance is offered via a feedback rule construction mechanism. To our knowledge, it is the first attempt in the literature to develop this kind of diagnosis and assessment tool. Experimental results show that the proposed learning aid mechanism can effectively assist low ability learners



in making progress during the continuous development of the assigned projects, and the assessment module is confirmed to be capable of evaluating the quality of learner's work correctly as well.

The remainder of the paper is organized as follows. The related work is given in next section. The overall architecture of the e-learning platform, the diagnosis and feedback module used in the platform, the assessment module employed for quality evaluation of learners' work, and the experimental results are discussed in respective sections. Finally, conclusions and the future work are provided in the last section.

## Related work

In the past few years, developing useful learning diagnosis and assessment systems using machine learning techniques has become a hot research topic in the literature (Raineri *et al.*, 1997; Smaill, 2005; McGourty, 2000; Zhang *et al.*, 2001; Cheung *et al.*, 2003; Cheng *et al.*, 2005; Tsaganou *et al.*, 2003; Lo *et al.*, 2004; Depradine *et al.*, 2003; Guzmán *et al.*, 2005). As the Internet gains wide popularity around the world, e-learning is taken by the learners as an important study aid. In order to help teachers easily analyze students' portfolios in an intelligent tutoring system, many researchers try to extract some useful information from the portfolios and reflect the degree of students' participation in the curriculum activity. The intelligence of intelligent tutoring systems is seen through the way these intelligent systems adapt themselves to each individual student, such as speed of learning, specific areas in which the student excels as well as falls behind, and rate of learning as more knowledge is learned. The interactions among the students and the intelligent tutoring system realized by Piramuthu (Piramuthu, 2005) include instructing, evaluating feedback from students, learning the characteristics of students, tailoring instructions as per the characteristics and feedback received, and being able to adjust to variable student learning rates.

Intelligent educational systems benefit enormously from their ability of providing adaptivity and customization, thus maximizing system personalization (Aroyo *et al.*, 2007). It was reported by Raineri *et al.* (1997) that prompt feedbacks given by an intelligent tutoring system can effectively reduce the sense of frustration felt by the students. The main benefits of the implementation of web-based learning and assessment tool have been described in (Smaill, 2005; McGourty, 2000). It was observed that students have been highly motivated, have received regular and well-timed feedbacks on their progress, and have promoted their achievement levels; instructors have been able to manage workloads while maintaining effective teaching and assessment practices. Using hints and feedbacks (Guzmán *et al.*, 2005), students can participate in an active learning process and can contribute to improving students' knowledge and detecting some possible misconceptions.

The natural language processing (NLP) is a way of computer processing of human language, and the computer returns a correct response that people expect. The NLP can be applied to language-based fields, such as speech recognition, dialogue system, information retrieval and knowledge representation (Alshawi, 1992). A lot of literature studies and optimization techniques have been also successfully applied to NLP, such as neural networks (Miikkulainen, 1993; Schmid, 1994), rule-based reasoning (Yang, 2000), pattern matching (Navarro & Raffinot, 2002) and text mining (Salton, 1989). The NLP represents the meaning of each sentence based on the built-in grammar and semantic database, such as ontology or rule-based database. However, building one exhaustive grammar and semantic database is a difficult task, especially for some specific application domains, such as the interpretation of program comments given by the learners as required in this work.

Recently, text mining techniques have been widely applied in many research domains, such as semantic web mining (Eikvil, 1999) and foreign exchange prediction (Permuntilleke & Wong, 2002). Text mining involves looking for patterns in natural language text and may be defined as the process of analyzing text to extract information. For instance, sentence similarity is used as a criterion to extract unseen knowledge from textual database. It has been proven that it is an effective way to retrieve information from short text. The text mining techniques have been employed in the implementation of intelligent e-learning platforms as well. Yuan & Chen (2007) integrated unique association thinking of humans with an automated decision agent grounded on the three association capabilities of human thinking, including similarity, contiguity and contrast, such that the students can participate actively in brainstorming. Pantic *et al.* (2005) employed an intelligent agent with the semantic network concepts to teach introductory artificial intelligence and relevant information can be monitored, filtered, and retrieved from the Web.

Text mining involves looking for patterns in natural language text and may be defined as the process of analyzing text to extract information. For instance, sentence similarity is used as a criterion to extract unseen knowledge form textual database. It has been proven effectively way to retrieved information from short text. Traditional techniques for detecting similarity between long texts documents have focused on analyzing shared words (Salton, 1989). However, word co-occurrence in sentence similarity may be scarce or even nonexistent mainly because of the inherent flexibility of natural language, which enables people to express similar meanings using quite different sentences in terms of structure and content. To tackle this challenge, the focus of this paper is primarily on computing the similarity between very short texts, such as program comments given by the learners during the development of the programs.

Although sentence similarity is increasingly in demand from a variety of applications, the adaptation of available measures to computing sentence similarity has several major drawbacks. Firstly, a sentence is represented in a very high-dimensional space with hundreds or thousands of dimensions (Salton, 1989; Landauer *et al*, 1998). This results in a very sparse sentence vector which is consequently computationally inefficient. High dimensionality and high sparsity can also lead to unacceptable performance in similarity computation (Burgess *et al*, 1998). Moreover, once the similarity computing method is designed for a specific application domain, it cannot be adapted easily to other domains. The lack of adaptability does not correspond to human language usage because sentence meaning might vary from domain to domain to certain extent. For instance, the well-known text analyzing tool in the literature, latent semantic analysis (LSA), attempts to represent all the semantic information in the document by a single low dimensional vector. Nevertheless, the variety of the topic, the complexity of text structure, plus the flexibility of personal expression and the ambiguity of natural language, make it impossible to realize this goal (Zhang & Rudnicky, 2002).

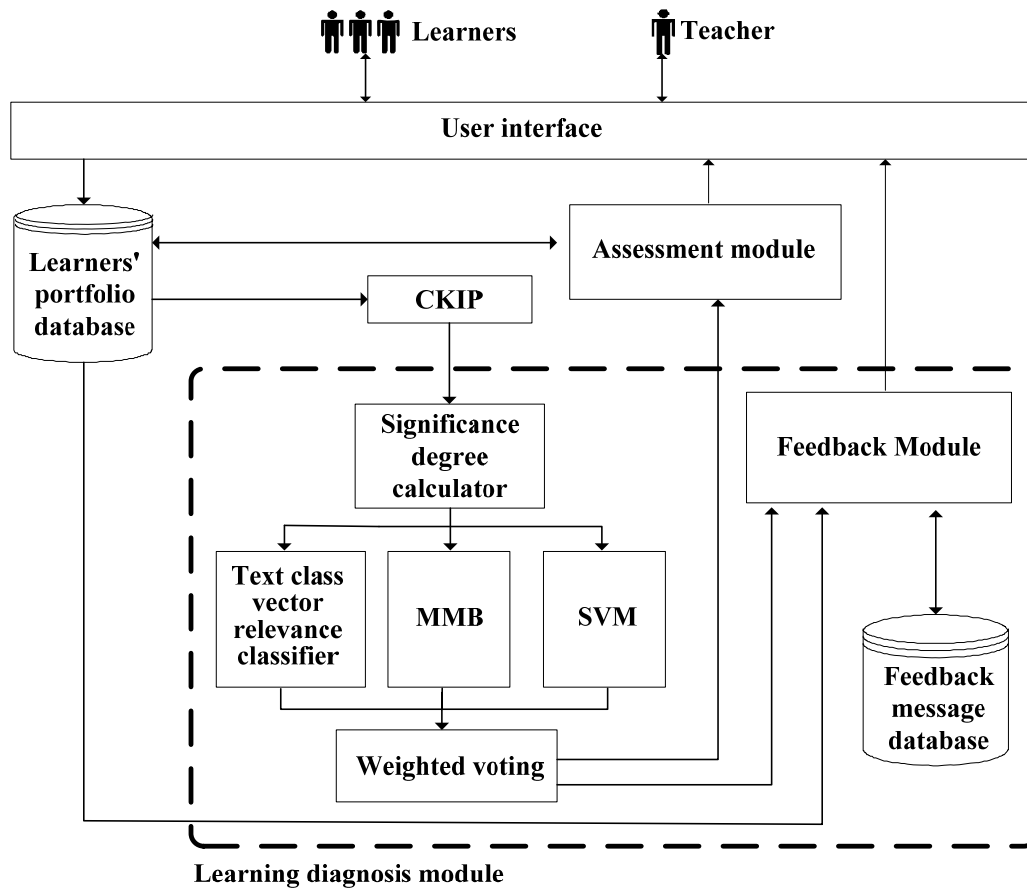


Figure 1. Architecture of the e-learning platform for a programming course

## Architecture of the e-learning platform

Figure 1 shows the architecture of the e-learning platform used in this work. As seen from the figure, the e-learning platform is employed to assist the teacher in teaching at programming courses for undergraduate students. There are three main components in the e-learning platform, which includes a diagnosis module, an assessment module, and the user interface module. The inputs to both the diagnosis and the assessment modules are obtained from the learners' portfolios updated during online learning activities in our e-learning platform as shown in Figure 2. The learners' portfolios includes each learner's ability estimation given by the teacher based on the historical learning records of each learner, the response time of each learner and the source code along with the comments posted by each learner on the e-learning platform. The diagnosis module monitors the learners' learning status and gives timely support to the students that need help. The assessment module evaluates the students' learning outcomes using the learning portfolios and gives the final reports to the teacher and the students.

The assessment module as shown in Figure 1 is employed to give the quality evaluation of the source code developed by the learners. The measurement given by the assessment module is based on the execution result of the learner's program, the comments given by the learner, the program elements, and the structure organized by the learner. The measurement allows the learners to have the opportunity to improve their code quality before submitting the final work to the teacher.

**資料結構使用C** [使用者: 張書銘] [回首頁](#)

請描述九九乘法表題目的特性與解題想法: 如描述資料規則性, 程式的流程控制與語法描述

**Detailed descriptions of this assigned project**

1. 假設變數 X、Y、Z  
2. 用雙重 for 迴圈  
3. 限制變數 X、Y (X、Y=1, X、Y<=9) 遞增直到 X、Y 等於 10 跳出 for 迴圈  
4. Z=X\*Y, 計算出結果,  
5. print 出九九乘法表

請將程式碼寫在下列區塊 [送出編譯](#)

```
#include<stdio.h>

int main(void)
{
    int i,j,x;          //宣告 i,j,x 為整數
    for(i=1;i<=9;i++)   //for 迴圈 i, i<=9, i 遞增
    {
        for(j=1;j<=9;j++) //for 迴圈 j, j<=9, j 遞增
        {
            x = i*j;      //算式

            printf("%d x %d = %d ~ \n\n", i,j,x); //顯示結果
            printf("\n");  //空行
        }
    }
    return 0;           //傳回 0
}
//end
```

**Program source code with appropriate comments**

**Moodle E-learning Platform for Programming Courses**

**紀錄**

一. 編譯錯誤紀錄:

二. 程式執行時間: 1.139秒

三. 記憶體的大小: 9k

四. 程式執行結果:

**Effectiveness of executing program**

**Executing result of program**

```
1 1 x 1 = 1 ~ 1 x 1 = 1
2 1 x 2 = 2 ~ 1 x 2 = 2
3 1 x 3 = 3 ~ 1 x 3 = 3
4 1 x 4 = 4 ~ 1 x 4 = 4
5 1 x 5 = 5 ~ 1 x 5 = 5
6 1 x 6 = 6 ~ 1 x 6 = 6
7 1 x 7 = 7 ~ 1 x 7 = 7
8 1 x 8 = 8 ~ 1 x 8 = 8
9 1 x 9 = 9 ~ 1 x 9 = 9
10 2 x 1 = 2 ~ 2 x 1 = 2
11 2 x 2 = 4 ~ 2 x 2 = 4
12 2 x 3 = 6 ~ 2 x 3 = 6
13 2 x 4 = 8 ~ 2 x 4 = 8
14 2 x 5 = 10 ~ 2 x 5 = 10
15 2 x 6 = 12 ~ 2 x 6 = 12
16 2 x 7 = 14 ~ 2 x 7 = 14
17 2 x 8 = 16 ~ 2 x 8 = 16
18 2 x 9 = 18 ~ 2 x 9 = 18
19 3 x 1 = 3 ~ 3 x 1 = 3
20 3 x 2 = 6 ~ 3 x 2 = 6
21 3 x 3 = 9 ~ 3 x 3 = 9
22 3 x 4 = 12 ~ 3 x 4 = 12
23 3 x 5 = 15 ~ 3 x 5 = 15
24 3 x 6 = 18 ~ 3 x 6 = 18
25 3 x 7 = 21 ~ 3 x 7 = 21
26 3 x 8 = 24 ~ 3 x 8 = 24
27 3 x 9 = 27 ~ 3 x 9 = 27
28 4 x 1 = 4 ~ 4 x 1 = 4
29 4 x 2 = 8 ~ 4 x 2 = 8
30 4 x 3 = 12 ~ 4 x 3 = 12
31 4 x 4 = 16 ~ 4 x 4 = 16
32 4 x 5 = 20 ~ 4 x 5 = 20
33 4 x 6 = 24 ~ 4 x 6 = 24
34 4 x 7 = 28 ~ 4 x 7 = 28
35 4 x 8 = 32 ~ 4 x 8 = 32
36 4 x 9 = 36 ~ 4 x 9 = 36
37 5 x 1 = 5 ~ 5 x 1 = 5
38 5 x 2 = 10 ~ 5 x 2 = 10
39 5 x 3 = 15 ~ 5 x 3 = 15
40 5 x 4 = 20 ~ 5 x 4 = 20
41 5 x 5 = 25 ~ 5 x 5 = 25
42 5 x 6 = 30 ~ 5 x 6 = 30
43 5 x 7 = 35 ~ 5 x 7 = 35
44 5 x 8 = 40 ~ 5 x 8 = 40
45 5 x 9 = 45 ~ 5 x 9 = 45
46 6 x 1 = 6 ~ 6 x 1 = 6
47 6 x 2 = 12 ~ 6 x 2 = 12
48 6 x 3 = 18 ~ 6 x 3 = 18
49 6 x 4 = 24 ~ 6 x 4 = 24
50 6 x 5 = 30 ~ 6 x 5 = 30
51 6 x 6 = 36 ~ 6 x 6 = 36
52 6 x 7 = 42 ~ 6 x 7 = 42
53 6 x 8 = 48 ~ 6 x 8 = 48
54 6 x 9 = 54 ~ 6 x 9 = 54
55 7 x 1 = 7 ~ 7 x 1 = 7
56 7 x 2 = 14 ~ 7 x 2 = 14
57 7 x 3 = 21 ~ 7 x 3 = 21
58 7 x 4 = 28 ~ 7 x 4 = 28
59 7 x 5 = 35 ~ 7 x 5 = 35
60 7 x 6 = 42 ~ 7 x 6 = 42
61 7 x 7 = 49 ~ 7 x 7 = 49
62 7 x 8 = 56 ~ 7 x 8 = 56
63 7 x 9 = 63 ~ 7 x 9 = 63
64 8 x 1 = 8 ~ 8 x 1 = 8
65 8 x 2 = 16 ~ 8 x 2 = 16
66 8 x 3 = 24 ~ 8 x 3 = 24
67 8 x 4 = 32 ~ 8 x 4 = 32
68 8 x 5 = 40 ~ 8 x 5 = 40
69 8 x 6 = 48 ~ 8 x 6 = 48
70 8 x 7 = 56 ~ 8 x 7 = 56
71 8 x 8 = 64 ~ 8 x 8 = 64
72 8 x 9 = 72 ~ 8 x 9 = 72
73 9 x 1 = 9 ~ 9 x 1 = 9
74 9 x 2 = 18 ~ 9 x 2 = 18
75 9 x 3 = 27 ~ 9 x 3 = 27
76 9 x 4 = 36 ~ 9 x 4 = 36
77 9 x 5 = 45 ~ 9 x 5 = 45
78 9 x 6 = 54 ~ 9 x 6 = 54
79 9 x 7 = 63 ~ 9 x 7 = 63
80 9 x 8 = 72 ~ 9 x 8 = 72
81 9 x 9 = 81 ~ 9 x 9 = 81
```

Figure 2. Screenshot of the e-learning platform for programming courses

## The intelligent learning diagnosis module

There are three major components in the learning diagnosis module as shown inside the blue dotted frame in Fig. 1. They are significance degree calculator, composite classifier, and feedback module. The three classifiers are adopted in this work to assist in computing the relevance degree between the learner's comments and the expected approaches. Each learner's source code comments given in Chinese are first segmented into separate keywords by a so-called Chinese Knowledge and Information Processing (CKIP) system (Chen & Liu, 1992; Chen & Bai, 1998; Chen & Ma, 2002; Ma & Chen, 2003). The significance degree calculator is used to derive numeric values that give the validity measure of learners' planning for their solutions based on the verbs and the nouns collected from the learner's comments. The derived numeric values are then fed into a composite classifier to determine if the learners go in the wrong direction. The feedback module is requested to give some feedback messages to the learners and the teacher based on a set of feedback rules built in the feedback message database if needed. Notably, the learning diagnosis module proposed in this work is not linked to any specific language or characters, because the inputs to the module, which are the verbs and the nouns extracted from the comments, are converted into the form of a numeric array for the ease of the computation.

The three base classifiers employed in the composite classifier are a text class vector relevance classifier, support vector machines (SVM), and a multimembership Bayesian classifier (MMB). The three base classifier models are trained on the collected sample source code comments and their precision is recorded. The precision is then used as the weight for each individual classifier. Thus, greater the precision of any classifier, the larger will be the weight associated with it. Notably, the SVM is chosen as one of the base classifiers in this work because they have been successfully employed in a wide range of real-world classification applications, such as handwritten digit recognition (Burges, 1998), speaker identification (Wan & Campbell 2000), and text categorization (Joachims, 1999). The MMB is employed here owing to its capability of efficiently modeling the uncertainty inherent to human reasoning in pattern recognition and machine intelligence fields. (Chang *et al.*, 1994; Lee *et al.*, 1991). Comparison to the Bayesian method, which can only identify a single category of content and the category must be assigned by the operators in advance, the MMB approach not only accurately identifies multi-category of the content in time but also has no restriction on collecting the training sets.

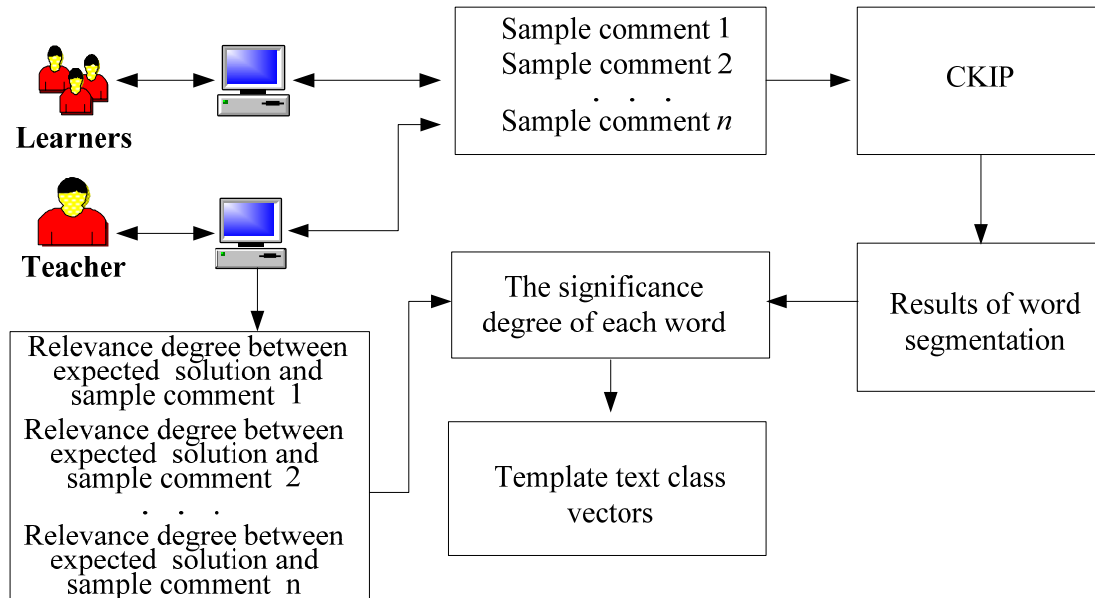


Figure 3. Text class vector construction process

### Significance degree calculator

Figure 3 shows the process of template text class vector construction in the significance degree calculator module. In this mechanism, the instructor first collects the sample comments and assigns the relevance degree for each sample

comment before the operation of the learning platform. The CKIP system developed by Academia Sinica in Taiwan is then employed to separate the Chinese words in each collected sample comment. The combination of word segmentation result and the value of relevance degree generated for each sample comment are used to measure the relevance between the comments given by the learner and the expected comments accordingly. Based on the relevance degree of each sample comment assigned by the instructor, the system is able to build a positive and a negative template class vector which stand for the relevance and irrelevance of the sample comments to the expected comments, respectively. The two template class vectors assist the learning platform in assigning a suitable degree value for each newly posted comment by students as an effective measurement tool for each student's learning status during the progress of the problem.

A summary of template text class vector construction algorithm is given below:

- (i) It is assumed that there are  $n$  sample comments collected by the instructor beforehand. They are fed into the CKIP system to extract each individual word in a Chinese sentence. Only verbs and nouns are kept for further processing. The word segmentation result for the  $i$ th sample comment is  $[W_{i,1}, W_{i,2}, \dots, W_{i,kl}]^T$ , where  $kl$  denotes the index for the last extracted word in the  $i$ th sample comment.
- (ii) According to the judgment of the instructor on the relevance degree between the sample comments and the proper approaches expected by the instructor, words in each sample comment can be placed in either positive class which is relevant to the expected solution, or negative class which is regarded as being an irrelevant or improper comment. The template text class vector for both classes can be expressed as:  $[P_1, P_2, P_3, \dots, P_{pos}]$  and  $[N_1, N_2, N_3, \dots, N_{neg}]$  in mathematics form, where the indices  $pos$  and  $neg$  denote the counts of the words collected in positive class and negative class, respectively.
- (iii) The significance degree of positive class and negative class for word  $P_i$  in each template text class vector can be expressed by the following equations:

$$sd_{P_i, pos} = \frac{pf_{P_i, pos} + \kappa}{pf_{P_i, neg} + \kappa} \cdot IDF_{P_i} \quad (1)$$

$$sd_{P_i, neg} = \frac{pf_{P_i, neg} + \kappa}{pf_{P_i, pos} + \kappa} \cdot IDF_{P_i}, \quad (2)$$

where  $pf_{P_i, pos}$  and  $pf_{P_i, neg}$  represent the normalized occurrence frequency of word  $P_i$  in positive class and negative class, respectively. A positive constant  $\kappa$  is added to prevent from dividing by zero in both equations. The value of  $\kappa$  is determined by the experiments.  $IDF_{P_i}$  is the so-called inverse document frequency (IDF), which is defined as:

$$IDF_{P_i} = \log \frac{N + \eta}{PN_{P_i} + \eta} \quad (3)$$

where  $N$  denotes the total number of comments,  $PN_{P_i}$  is the number of the comments that contains word  $P_i$ , and  $\eta$  represents a positive constant.

An example list of significance degree generation for extracted keywords is shown in Figure 4. Based on the occurrence frequency of each keyword for both classes in the sample comments, the normalized occurrence frequency of each keyword in both classes is first computed, as illustrated at Columns 4 and 5. Then the significance degrees of each keyword in both classes, as respectively given at Columns 7 and 8, can be derived by using Eqs. (1), (2), and (3).

- (iv) The two template text class vectors for both classes are formed by:

$$\vec{v}_{pos} = [sd_{P_1, pos}, sd_{P_2, pos}, \dots, sd_{P_{pos}, pos}], \quad (4)$$

$$\vec{v}_{neg} = [sd_{P_1, neg}, sd_{P_2, neg}, \dots, sd_{P_{neg}, neg}]. \quad (5)$$

Eqs. (4) and (5) can be used to compare with the word segmentation result of the  $i$ th comment given by the learner in the classification process.

The test text class vectors of positive class and negative class for the  $i$ th comment issued by the learner are then expressed by:

$$\vec{t}_{i,pos} = [sd_{p_1,pos}, sd_{p_2,pos}, \dots, sd_{p_{pos},pos}]$$

where  $sd_{p_j,pos} = 0$ , if  $p_j \notin \text{comment } i$ ,  $1 \leq j \leq pos$ ,

$$\vec{t}_{i,neg} = [sd_{p_1,neg}, sd_{p_2,neg}, \dots, sd_{p_{neg},neg}]$$

where  $sd_{p_j,neg} = 0$ , if  $p_j \notin \text{comment } i$ ,  $1 \leq j \leq neg$ .

The two test text vectors of each comment are used as the inputs for each individual classifier as shown in Figure 1 to determine the relevance degree to the expected comments.

詞	count_pos	count_neg	pf_pos	pf_neg	IDF	sd_pos	sd_neg
迴圈	30	0	0.0389	0	1.76451	1.833167797	1.698423316
for迴圈	30	0	0.0389	0	1.76451	1.833167797	1.698423316
印出	6	0.0389	0.048	1.587579	1.573809757	1.601468949	1.698423316
乘法表	149	29	0.1933	0.232	5.198497	3.833612309	3.50728997
輸出	30	6	0.0389	0.048	1.587579	1.573809757	1.601468949
列印	30	6	0.0389	0.048	1.587579	1.573809757	1.601468949
雙層迴圈	30	0	0.0389	0	1.76451	1.833167797	1.698423316
雙層while迴圈	30	0	0.0389	0	1.76451	1.833167797	1.698423316
while迴圈	21	0	0.0272	0	2.107455	2.107455	2.051575101
遞迴	30	0	0.0389	0	1.76451	1.833167797	1.698423316
陣列	30	0	0.0389	0	1.76451	1.833167797	1.698423316
堆疊	0	51	0	0.04	3.406738	3.275709194	3.543007065
佇列	0	51	0	0.04	3.406738	3.275709194	3.543007065
鏈結	0	51	0	0.04	3.406738	3.275709194	3.543007065
串列	0	51	0	0.04	3.406738	3.275709194	3.543007065
排序	0	51	0	0.04	3.406738	3.275709194	3.543007065
圖形	0	51	0	0.04	3.406738	3.275709194	3.543007065
堆積	0	51	0	0.04	3.406738	3.275709194	3.543007065
乘法表	149	29	0.1933	0.232	5.198497	3.833612309	3.50728997

Figure 4. Example list of significance degree generation for extracted keywords

### Text class vector relevance classifier

Figure 5 illustrates the architecture of text vector classification. The relevance degree of the expected comments and the comments issued by the learner is computed by the following formulae:

$$r_{pos} = \frac{\vec{v}_{pos} \cdot \vec{t}_{i,pos}}{\|\vec{v}_{pos}\| \|\vec{t}_{i,pos}\|}$$

$$r_{neg} = \frac{\vec{v}_{neg} \cdot \vec{t}_{i,neg}}{\|\vec{v}_{neg}\| \|\vec{t}_{i,neg}\|}$$

where  $\vec{v}_{pos}$  and  $\vec{v}_{neg}$  represent the template text class vectors for positive and negative classes, respectively, and the  $\vec{t}_{i,pos}$  and  $\vec{t}_{i,neg}$  denote the text class vectors of the  $i$ th comment issued by the learners for positive and negative classes, respectively.

The relevance degree computed by Eqs. (8) and (9) determines if the posted comments are relevant to the expected solution. The comments are predicted as being relevant to the expected solution if  $r_{pos}$  is larger than  $r_{neg}$  in Eqs. (8) and (9).

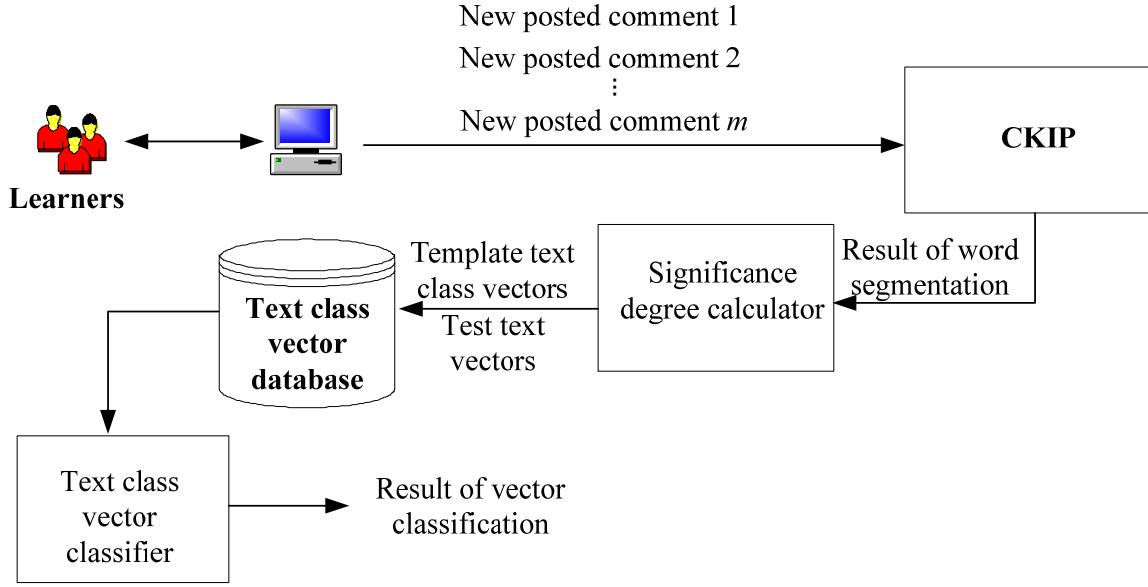


Figure 5. Process of text class vector classification

### SVM classifier

Recently, support vector machines (SVM) have gaining popularity due to its numerous attractive features and eminent empirical performance (Vapnik, 1995; Vapnik, 1998; Cherkassky & Mulier, 1999; Chang & Lin, 2001; Taha, 1997; Burges, 1998). To solve a nonlinear regression or functional approximation problem, the SVM nonlinearly map the input space into a high-dimensional feature space via a suitable kernel representation, such as polynomials and radial basis functions with Gaussian kernels. This approach is expected to construct a linear regression hyperplane in the feature space, which is nonlinear in the original input space.

In this work, the sample comments are used to train the SVM as shown in Figure 6 (a). The relevance degree to the expected solution for each segmented keyword in the sample comments can be expressed as:

$$rd_{w_{i,j}} = \frac{pf_{w_{i,j},pos} + \eta}{pf_{w_{i,j},neg} + \eta} \cdot IDF_{w_{i,j}}, \quad (21)$$

where word  $w_{i,j}$  denotes the extracted keyword,  $pf_{w_{i,j},pos}$  and  $pf_{w_{i,j},neg}$  represent the occurrence frequency of word  $w_{i,j}$  in the collected relevant comments and irrelevant comments, respectively.  $\eta$  is a positive constant, and  $IDF_{w_{i,j}}$  is given by Eq. (3).

The classification process for the posted comments is illustrated in Figure 6 (b), where the average of the classification result for the  $m$  posted comments of the learner is taken as the output of the SVM classifier.

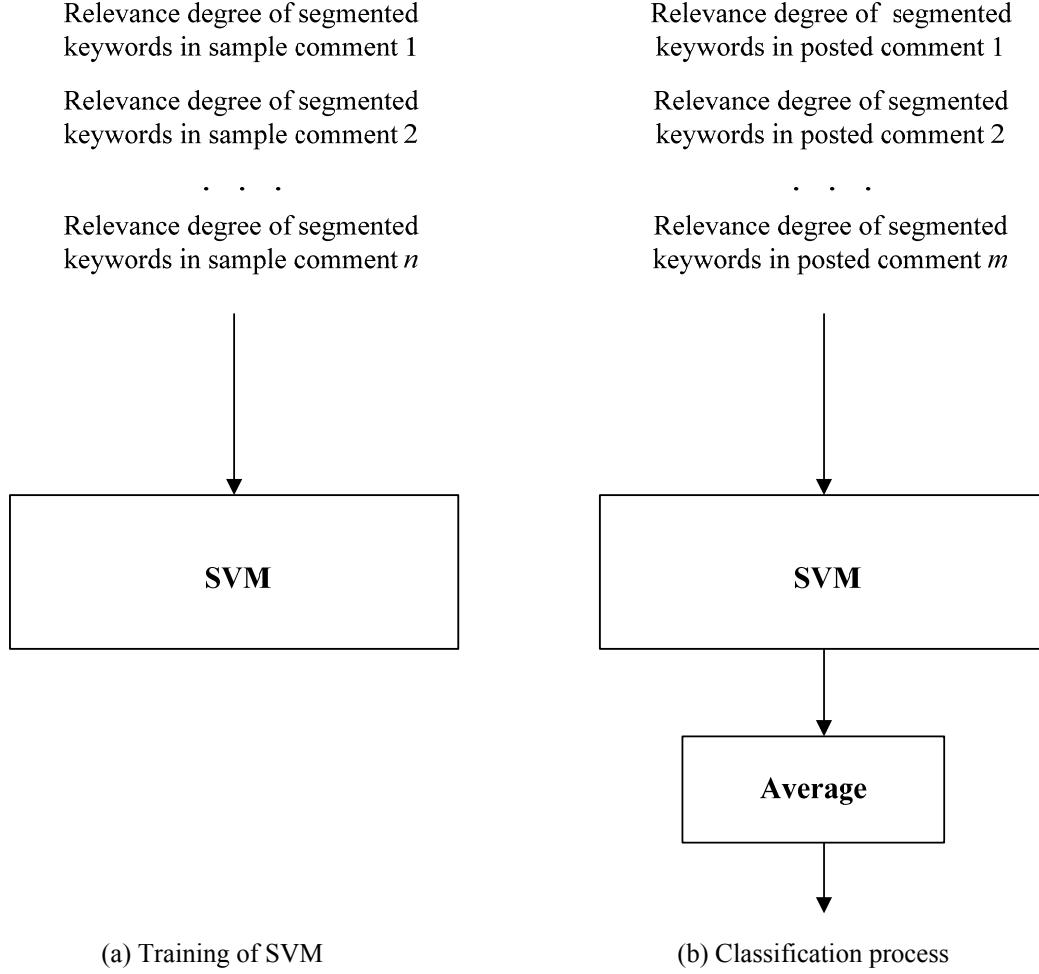


Figure 6. The Training and classification process of SVM

### Multimembership Bayesian classifier

Multimembership Bayesian (MMB) algorithm was adopted as the core module in the design of a medical diagnostic expert system and a web site classifier in the literature (Lo *et al*, 2005; Lo *et al*, 2006; Chang *et al*, 1994), and the reported experimental results were impressive. This work thus employs MMB as one of the three components in the composite classifier used in the intelligent diagnosis module.

The heart of the MMB classifier is its inference engine, which can be formulated as follows:

$$P(C_i | T_1, T_2, \dots, T_k) = \frac{P(C_i) \cdot P(T_1 | C_i) \cdot \dots \cdot P(T_k | C_i)}{P(C_i) \cdot P(T_1 | C_i) \cdot \dots \cdot P(T_k | C_i) + (1 - P(C_i)) \cdot P(T_1 | \bar{C}_i) \cdot \dots \cdot P(T_k | \bar{C}_i)} \quad (22)$$

where  $P(C_i)$  is the prior probability of the category  $C_i$ ,  $P(C_i | T_1, T_2, \dots, T_k)$  represents the posterior probability for category pattern  $C_i$  in related to the selected feature set of  $T_1, T_2, \dots, T_k$ , and  $P(T_j | C_i)$  denotes the conditional probability for feature  $T_j$  given category pattern  $C_i$  is present.

In this work, the relevance degree to the expected solution for each segmented keyword in the sample comments as given in Eqs. (1) and (2) is used to determine the feature set collected by the MMB inference engine, and equal number of the features are obtained from each of the two classes as shown in Figure 7 Then the corresponding test



text class vectors of positive and negative classes for the comments issued by the learner as given in Eqs. (6) and (7) are fed into the MMB inference engine to apply the MMB algorithm to obtain the classification result.

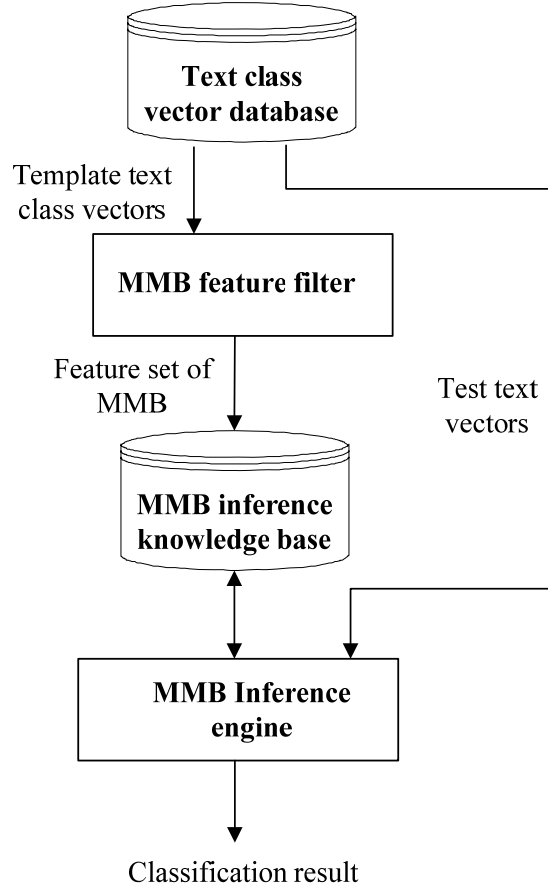


Figure 7. The MMB inference process

### Feedback module

Figure 8 shows the architecture of the feedback module, which is composed of a feedback rule constructor and a feedback message dispatcher. The feedback rule constructor allows the instructors to mark the sample comments that are irrelevant to the expected solution and create a corresponding rule. The consequent of each rule is also provided by the instructor as a way to give the learner a hint or appropriate suggestion. The feedback message dispatcher is activated to send the appropriate feedback message to the learner if the final decision from the composite classifier determines that the comments issued by learner are recognized as irrelevant to expected solution, and the required elements or structures designated by the teacher are missing in the learner's source code.

Whenever the comment posted by the learner is deemed as improper approach for developing the problem, all the sample comments for the negative class are examined and the one most similar to the posted comment will trigger the corresponding feedback rule to issue an appropriate diagnosis message back to the learner. The similarity between the sample and the posted comments is given by:

$$s = \frac{\vec{p} \cdot \vec{c}_i}{|\vec{p}| |\vec{c}_i|}, \quad (23)$$

$$\vec{c}_i = [sd_{1,i}, sd_{2,i}, \dots, sd_{k,i}], \quad (24)$$

$$\vec{p} = [sd_1, sd_2, \dots, sd_m], \quad (25)$$

where  $\vec{c}_i$  represents the text vector of the  $i$ th sample comment,  $\vec{p}$  denotes the text vector of the comment issued by the learners,  $k$  and  $m$  stand for the number of segmented keyword in the  $i$ th sample comment and the posted comment, respectively. Notably, the value of each element in the two text vectors is identical to the significance degree computed by Eq. (2).

Figure 8. Architecture of the feedback module

A sample feedback message generated by the feedback module is given in *Figure 9*. In this example, a learner was requested to build a 9×9 multiplication table by using PHP program language. The learner’s comments mentioned that each element in the 9×9 multiplication table is printed out one by one in his code. The composite classifier first determined that the learner’s comment is possibly irrelevant to the expected solution. Then the feedback message dispatcher is asked to verify that the required iteration structure is indeed missing in the learner’s source code. Accordingly, some warning message like “nested for/while iteration structure is suggested” was issued to the learner.



To justify the accuracy of the feedback messages generated by the feedback module, the tool further will pop a message window as shown in Figure 10 to ask the learner whether the issued feedback message is appropriate. The teacher will receive a message from the feedback module indicating the existence of some inappropriate feedbacks reported from the students. Figure 11 gives an example of an inappropriate feedback message that was reported by the students. Notably, the inappropriate feedback message is marked in red to allow the teacher to easily locate the inappropriate feedback messages that were reflected by the students.

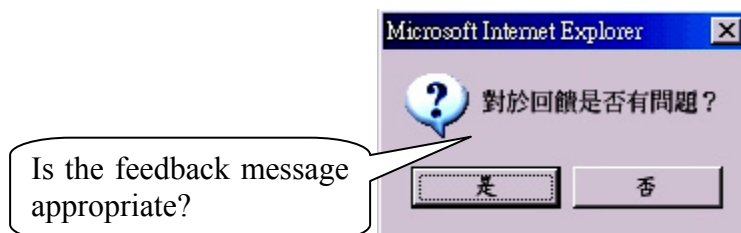


Figure 10. Confirmation on the correctness of feedback message



- Software verification and validation: All analysis and design software requirements and software design have been verified and validated for correct operation using testing, observation, or inspection techniques; Relevant abnormal conditions have been evaluated for mitigating unintended functions through testing, observation, or inspection techniques.
- Software quality assurance: SQA activities and software practices for requirements management, software design, software configuration management, procurement controls, V&V (including reviews and testing), and documentation have been evaluated and established at the appropriate level for proper applicability to the analysis and design software under assessment.

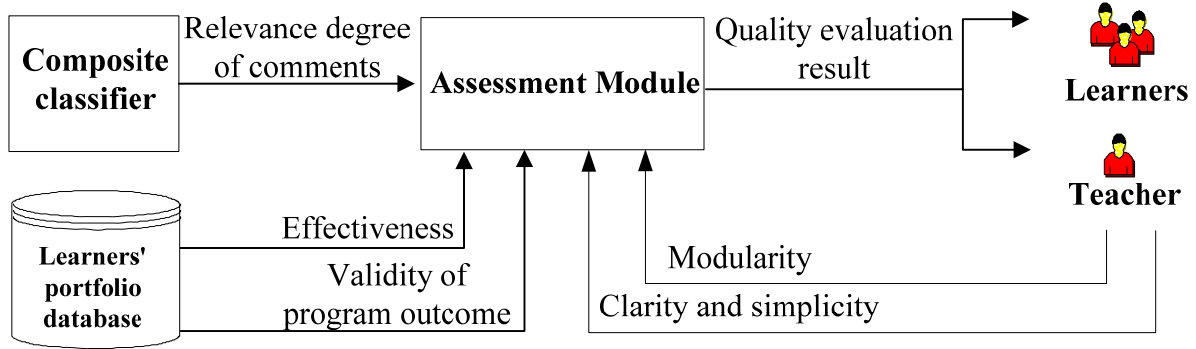


Figure 12. Assessment tool for evaluating the quality of learner's source code

Figure 12 illustrates the architecture of program source code assessment module used in the e-learning platform. After the learners submit their program source codes to the learning platform, the built-in assessment module can assist the teacher in evaluating the quality of each learner's work based on the following criteria proposed by Almutka et al (2004):

- Modularity: including various class implementation issues, such as declaring explicit constructors and assignment operator, explicit calls for inherited assignment operators. Encapsulation issues, such as class inheritance, public member variables belong to this category.
- Clarity and simplicity: such as the average length of functions, blocks with braces, and usage of short-circuit statements.
- Effectiveness, such as small variable scopes.
- Documentation: whether meaningful comments are given.
- Validity of program outcome: including correctness and rationality of the execution output.

Notably, the motivation of choosing this particular set of criteria is to assess the practical coding and design conventions of the students' programs because it was essential to have more knowledge of the program structure than merely to collect statistics from the source code. Several features that are already analyzed by PHP or C compilers when given full warning options, such as unused variables or parameters, were decided not to be implemented in our assessment module.

In this work, one or more sample program source codes are offered by the teacher for each project, and the given sample programs are used as a reference to grade the learner's work. The first two criteria, including the modularity and the clarity/simplicity of program source code, are explicitly provided by the teacher at the end of the project development. The degree of effectiveness can be obtained based on the learner's program executing time and memory usage. The validity of documentation is assessed by the composite classifier employed in the diagnosis module, in which the relevance degree of learner's concept to the solution of the assigned problem is compared. The learner's program outcomes, including correctness and rationality of the execution output, are contrasted with the expected sample program outcomes provided by the teacher. Notably, the teacher is allowed to inspect and amend the grading policy of the criteria established by the assessment module.

After the assessment module collects the grading results for the five criteria, the assessment module can then evaluate the overall quality of the learner's work according to the following formula:

$$Grade = \alpha \cdot m + \beta \cdot s + \gamma \cdot e + \delta \cdot r + \varepsilon \cdot v \quad (26)$$

where  $m$ ,  $s$ ,  $e$ ,  $r$  and  $v$  represent the degrees of modularity, clarity and simplicity, effectiveness, validity of documentation and program outcome, respectively.  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ , and  $\varepsilon$  are the weighting factors for the corresponding terms in Eq. (26). The weighting factors are determined by the instructor before the problems are assigned to the learners. Figure 13 illustrates an example of the quality evaluation results given by the assessment module on a one-to-one-hundred scale.

編號	使用者名稱	程式內容與註解	主題相關度	可模組性	清晰及簡潔度	有效度	正確性
28	王聰明	<pre> &lt;?php /*遞迴將數字格式化函式*/ function t(\$n) {     if (\$n &lt; 1000)         return \$n;     else     {         if (\$n%1000 == 0)             return t(floor(\$n/1000)).",000";         else             return t(floor(\$n/1000))."." . (\$n%1000);         }     } </pre>	78	78.124	70.531	77	72
31	王聰明	<pre> function money why if (\$n &lt; 1000) /* 終止條件 */     return \$n; else {     if (\$n%1000 == 0)         return money((int)(\$n/1000)).",000";     else if (\$n % 1000 &gt; 99)         return money((int)(\$n/1000))."." . (\$n%1000);     else if (\$n % 1000 &gt; 9)         return money((int)(\$n/1000))."." . (\$n%1000);     else         return money((int)(\$n/1000))."." . (\$n%1000); } </pre>	82.45	75.235	72.15	77	96

Figure 13. An example of the quality evaluation result given by the assessment module

## Experimental Results and Analyses

To verify the effectiveness of diagnosis and assessment modules proposed in this work, three introductory-undergraduate programming classes, including a PHP and two C programming courses, participated in the experiments. There were a total of 126 learners attending the programming classes, including 47 learners in the PHP programming class; 37 and 42 learners in two C programming classes, respectively. All the classes spanned 18 weeks. The teacher in experimental group adopts the proposed learning diagnosis and assessment tools as a teaching aid besides ordinary classroom teaching activities; whereas only ordinary classroom teaching activities were organized at control group. The students in the two classes are divided into three ability groups. In experimental group, there were a total of 79 learners, including 21 learners from the PHP programming class and 58 learners from the C programming class. Each learner in experimental group was asked to work on the assigned programming projects by using the learning platform at the end of the 16th weeks. Rest of the 47 learners, including 26 learners from the PHP programming class and 21 learners from the C programming class, are assigned in control group.

Each learner in experimental group is asked to give a brief description of his/her algorithm ahead of his/her source code. The program comments posted by the learners are separated from the source code and are fed into the significance degree calculator module to build the corresponding test text vector. The text vector for each comment was then fed into the composite classifier to determine the relevance degree between the source code comments and the expected solutions. The diagnosis tool generated feedback message to the learners whenever the comments and the source code constructed by the learners reveal that the learner deviate the expected approaches. The learning platform can automatically lower the difficulty level of the problems when the learners got stuck with the problem at the current difficulty level for a preset time limit.

Table 1 gives the comparison of the precision rates of three individual and composite classifiers. It can be observed that three individual classifiers and the composite classifier achieve satisfactory performance and the composite

classifier outperforms the three individual classifiers as expected. We can infer that the overwhelming majority of the relevance degree between the learner's comments and the expected approaches determined by the classifiers is trustworthy.

*Table 1.* Comparison of the precision rates of three individual and composite classifiers

Classifier	SVM	MMB	TVC	Weighted voting
Precision rate	95.24%	85.71%	92.86%	97.62%

Table 2 contrasts the quality evaluation results of the learner's work in the experimental PHP and C programming classes, graded by the teacher and the assessment module, in term of the following root mean square error:

$$E = \sqrt{\frac{\sum_{i=1}^n (g_{i,a} - g_{i,t})^2}{n}}, \quad (27)$$

where  $n$  represents the number of the learners,  $g_{i,a}$  denotes the  $i$ th student's grade determined by the assessment module, and  $g_{i,t}$  is the  $i$ th student's grade given by the teacher, both on a one-to-one-hundred scale.

*Table 2.* Comparison of the quality evaluation of the learner's program given by the teacher and the assessment module

Criteria	Modularity	Clarity & simplicity	Effectiveness	Relevance degree
Root mean square error (%)	3.875	5.912	2.03	2.995

As shown in Table 2, the root mean square error for the first four criteria as given in Section 5 is 3.703% on the average. It can be observed that the gap of the grading results for some criteria, such as clarity and simplicity, is not trivial between the teacher and the assessment module. The assessment module will inform the teacher and issue a request of reviewing and revising the grading policy of the learner's program for the corresponding criterion established in the assessment module for future reference.

Table 3 shows the comparison of problem accomplishment rates for the learners in the experimental between the beginning of 16th week and after. It can be seen that the learners improve their performance with the aid of the feedback module. Meanwhile, the learners in median and low ability groups that took use of the e-learning platform got improved significantly due to the effectiveness of the feedback messages sent by the diagnosis tool.

The comparison of project accomplishment rates for the learners between the beginning of 16th week and after in the three programming classes is given in Table 3. It can be seen that the learners in middle and low ability groups that took use of the e-learning platform got improved significantly owing to the effectiveness of the feedback message sent by the diagnosis module.

*Table 3.* Comparison of project accomplishment rates for the learners in three programming classes

Group Ability	Experimental group		Control group	
	Pre-test	Post-test	Pre-test	Post-test
High	73.7%(14/19)	100%(19/19)	69.2%(9/13)	76.9%(10/13)
Middle	42.4%(14/33)	69.7%(23/33)	26.7%(4/15)	33.3%(5/15)
Low	14.8%(4/27)	44.4%(12/27)	21.1%(4/19)	21.1%(4/19)

Tables 4 and 5 compare the students' achievement before and after they used diagnosis and assessment tools. The statistical results were obtained by running  $t$ -test with the SPSS software package. The average scores received by the learners in experimental group who were guided by the diagnosis and assessment tools are apparently better than that without the support. In addition, the evidence of high significant correlation between pre-guidance and post-guidance and the high significance level as given by the results of  $t$ -test as respectively illustrated in Table 4 and 5 indicate that the learners indeed made significant progress during the continuous development of the assigned programming projects with the assistance of the diagnosis and assessment tool.

Table 4. Paired samples statistics and correlations for the learners in experimental group

		N	Mean	Std. Deviation	Std. Error Mean	Correlation	Significance
<b>PHP programming class</b>	Pre-test	21	67.0476	13.25321	2.89209	0.775	0.000
	Post-test	21	76.6190	9.62536	2.10043		
<b>C programming classes</b>	Pre-test	58	62.8539	15.34193	2.91557	0.661	0.001
	Post-test	58	76.6538	10.50508	1.87963		

Table 5. Paired sample *t*-test for the learners in experimental group

	Mean	Standard deviation	Standard error mean	95% confidence interval of the difference		<i>t</i>	df	Significance (2-tailed)
				Lower	Upper			
<b>PHP programming class</b>	9.5714	11.4176	2.0448	5.7149	17.3831	5.017	20	0
<b>C programming classes</b>	-5.96205	12.42295	2.35275	-10.8064	-1.1176	-2.134	57	0.001

Table 6. Independent-samples *t*-test statistics for the learners in experimental and control groups

		N	Mean	Standard deviation	Standard error mean
<b>PHP programming class</b>	Experimental group	21	76.6190	9.69782	2.11624
	Control group	26	69.8077	11.17862	2.19231
<b>PHP programming class</b>	Experimental group	58	76.6538	10.50508	1.879625
	Control group	21	71.4286	17.89852	3.90578

Table 6 gives the *t*-test statistics for the students' grades in experimental and control groups. It can be observed that the mean of the students' grades in experimental group is higher than that for those in control group. That is, the students who received the help of the diagnosis and assessment tool achieve higher performance on average. The effectiveness of the proposed diagnosis and assessment tool is thus justified.

## Conclusions and Future Work

In this work, an intelligent learning diagnosis tool based on text mining and machine learning techniques is proposed. We use program comments and the source code, which are constructed by the learners during development of the assigned problems, to examine whether the learners plan wrong solutions. A feedback rule construction mechanism is used to issue feedback messages to the learners in case the diagnosis tool detects that the learners go in the wrong direction. An assessment module that evaluates quality of the learner's work is also proposed. Five measurement criteria including modularity, clarity and simplicity, effectiveness, documentation, and validity of program outcome are used as quality evaluation metrics. The instructor can either take the assessment results as a part of the learners' final grade, or use this assessment to uncover the learners that fall behind at the end of each learning activity stage.

Two introductory-undergraduate programming classes, including a PHP and a C programming course, participated in the experiments to verify the effectiveness of proposed diagnosis and assessment tools. The statistical results obtained by running *t*-test with the SPSS software package were given and analyzed. The experimental results exhibit that the proposed work is effective in assisting the middle and low learners in planning the solution for their programming assignment timely without the involvement of addition human resources anytime and anywhere.

In the future work, we plan to employ neuro-fuzzy expert systems technique to elaborate the design of the classifier and the feedback rule construction mechanism used in the diagnosis module. The reason of using the neuro-fuzzy expert system is that it can function more like human experts who explain the reasoning processes behind their recommendation. The fuzzy model successfully handles reasoning with imprecise information, and enables representation of student modeling in the linguistic form, which is the same way the human teachers do. Meanwhile, the neural networks embedded in the neuro-fuzzy expert systems, as observed in (Vrettaros *et al*, 2004; Huang, *et al*, 2003; Arriaga *et al*, 2002; Sevarac, 2006), can tune the membership functions and the rule parameters of the fuzzy model to improve the accuracy of the classification required in the learning diagnosis module. The last but not the least, more advanced text mining technique will be adopted in our future work to extract effective parameters from learners' learning portfolios as the inputs to the learning diagnosis tool.

Furthermore, we intend to incorporate other influencing factors, such as gender, heterogeneity of the class and grading schemes, along with the deployment of various collaborative learning mechanisms into the e-learning platform in the future work and investigate whether the ability of solution planning and bug removing for the middle and low ability learners can get further improved.

## Acknowledgment

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 95-2221-E-026-001.

## References

- Ala-Mutka, K., Uimonen, T., & Järvinen, H.-M. (2004). Supporting students in C++ Programming Courses with Automatic Program Style Assessment. *Journal of Information Technology Education*, 3, 245–262.
- Alshaw, H. (1992). The Core Language Engine. *ACL-MIT Press Series in Natural Language Processing*, Cambridge, MA: MIT Press.
- Aroyo, L., Graesser, A., & Johnson, L. (2007). Guest Editors' Introduction: Intelligent Educational Systems of the Present and Future. *IEEE Intelligent Systems*, 22 (4), 20–21.
- Arriaga, F. de, Laureano, A. L., Alami, M. El, & Arriaga, A. (2002). Some Applications of Fuzzy Logic to Intelligent Tutoring Systems, *Proc. International Conference on Technology and Education*, Badajoz, 1222-1227.
- Burgess, C., Livesay, K., & Lund, K. (1998). Explorations in context space: words, sentences, discourse. *Discourse Processes*, 25 (2-3), 211-257.
- Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*. 2 (2), 121–167.
- Chang, C.-C., & Lin, C.-J. (2001). *LIBSVM: a library for support vector machines*. Retrieved July 15, 2008 from <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Chang, L.-J., Evens, M., & Trace, D. A. (1994). A Knowledge Engineering System for MEDAS. In *Proceeding of 1994 IEEE Seventh Symposium on Computer-Based Medical Systems*, 6–11.
- Chen, K.J., & Bai, M.-H. (1998). Unknown word detection for Chinese by a corpus-based learning method. *International Journal of Computational linguistics and Chinese Language Processing*, 3 (1), 27-44.
- Chen, K.J., & Liu, S.H. (1992). Word identification for Mandarin Chinese sentences”, In *Proceedings of COLING 1992*, 101-107.
- Chen, K.J., & Ma, W.-Y. (2002). Unknown word extraction for Chinese documents. In *Proceedings of COLING 2002*, 169-175.
- Cheng, S.-Y., Lin, C.-S., Chen, H.-H., & Heh, J.-S. (2005). Learning and diagnosis of individual and class conceptual perspectives: an intelligent systems approach using clustering techniques. *Computers and Education*, 44 (3), 257-283.



- Cherkassky, V., & Mulier, F. (1999). Vapnik-chervonenkis (vc) learning theory and its applications. *IEEE Transactions on Neural Networks*, 10 (5), 985 – 987.
- Cheung, B., Hui, L., Zhang, J., & Yiu, S. M. (2003). SmartTutor: An intelligent tutoring system in web-based adult education. *The Journal of Systems and Software*, 68, 11-25.
- Depradine, C. (2003). Expert system for extracting syntactic information from Java code. *Expert Systems with Applications*, 25, 187-198.
- Eikvil, L. (1999). Information extraction from World Wide Web. *Technical Report 945*, Norwegian Computing Center.
- Fix, V., & Wiedenbeck, S. (1996). An intelligent tool to aid students in learning second and subsequent programming languages. *Computers & Education*, 27 (2), 71–83.
- Greyling J. H., Cilliers C. B., & Calitz A. P. (2006). B#: The Development and Assessment of an Iconic Programming Tool for Novice Programmers. In *Proc. of Information Technology Based Higher Education and Training*, 367-375.
- Guzmán, E., & Conejo, R. (2005). Self-Assessment in a Feasible, Adaptive Web-Based Testing System. *IEEE Transactions on Education*, 48 (4), 688-695.
- Huang, C.J., & Lai, W.K. (2003). Application of neuro-fuzzy technique to the bandwidth reservation for sectorized cellular communications. *Proceedings of 2003 IEEE Congress on Evolutionary Computation* (Vol. 2), 958-965.
- Hulls, C.C.W., Neale, A.J., Komalo, B.N., Petrov, V., & Brush, D.J (2005). Interactive Online Tutorial Assistance for a First Programming Course. *IEEE Transactions on Education*, 48 (4), 719–728.
- IEEE Std. 1012 (n.d.). *Standard for Software Verification and Validation*, retrieved July 15, 2008 from <http://ieeexplore.ieee.org/iel4/5672/15185/00702332.pdf>.
- Joachims, T. (1999). Transductive inference for text classification using support vector machines. *Proceedings of the Sixteenth International Conference on Machine Learning*, San Francisco, CA: Morgan Kaufmann, 200-209.
- Landauer, T.K., Foltz, P.W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, (2-3), 259-284.
- Lee, C.-Y., Evens, M., Carmony, L., Trace, D. A., & Naeymi-Rad, F. (1991). Recommending Tests in a Multimembership Bayesian Diagnostic Expert System. In *Proceedings of the Fourth Annual IEEE Symposium on Computer Based Medical Systems*, 28-35.
- Lo, J.-J., Wang, H.-M., & Yeh, S.-W. (2004). Effects of confidence scores and remedial instruction on prepositions learning in adaptive hypermedia. *Computers and Education*, 42, 45-63.
- Lo, S.-A., Lee, C.-Y. & Hsu C.-C. (2005). Using MMB Algorithm to Refine the Performance of Chinese Web Site Automatically Classified System. *Proceedings of the 2005 NCS International Conference*, retrieved July 15, 2008 from <http://hdl.handle.net/2377/2358>.
- Lo, S.-A., Lee, C.-Y., & Hsu C.-C. (2006). Automatically Classify Web Sites by Multimembership Bayesian Approach. In *Proceeding of Third International Conference on Information Technology: New Generations, ITNG 2006*, 580-583.
- Ma, W.-Y., & Chen, K.J. (2003). A bottom-up merging algorithm for Chinese unknown word extraction. In *Proceedings of ACL workshop on Chinese Language Processing 2003*, 31-38.
- McGourty, J. (2000) Using Multisource Feedback in the Classroom: A Computer-Based Approach. *IEEE Transactions on Education*, 43 (2), 120-124.
- Miikkulainen R. (1993), *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory*, Cambridge, MA: MIT Press.
- Mungunsukh, H., & Cheng, Z. (2002). An agent based programming language learning support system. In *Proceeding of the 2002 International Conference on Computers in Education*, 1, 148–152.
- Navarro, G., & Raffinot, M. (2002), *Flexible Pattern Matching in Strings—Practical On-Line Search Algorithms for Texts and Biological Sequences*, Cambridge, U.K.: Cambridge Univ. Press.

- Pantic, M., Zwitterloot, R., & Grootjans, R.J. (2005). Teaching Introductory Artificial Intelligence Using a Simple Agent Framework. *IEEE Transaction on Education*, 48 (3), 382–390.
- Permuntilleke, D., & Wong, R. K. (2002). Currency exchange rate forecasting from news headlines. *In Proceedings of the 13th Australian Database Conference*, Melbourne, Australia, 131-139.
- Piramuthu, S. (2005). Knowledge-Based Web-Enabled Agents and Intelligent Tutoring Systems. *IEEE Transaction on Education*, 48 (4), 750-756.
- Raineri D.M., Mehrtens B.G., & Hu bler A.W. (1997), CyberProf—An Intelligent Human-Computer Interface for Interactive Instruction on the World Wide Web. *Journal of Asynchronous Learning networks*, 1 (2), retrieved July 15, 2008 from [http://www.sloan-c.org/publications/jaln/v1n2/pdf/v1n2\\_raineri.pdf](http://www.sloan-c.org/publications/jaln/v1n2/pdf/v1n2_raineri.pdf).
- Salton, G. (1989). *Automatic Text Processing*, Addison-Wesley, Reading, Mass.
- Schmid, H. (1994). Part-of-speech tagging with neural networks. *In Proc. Int. Conf. Computational Linguistics*, 172–176.
- Sevarac, Z. (2006). Neuro Fuzzy Reasoner for Student Modeling. *Proceedings of the Sixth International Conference on Advanced Learning Technologies (ICALT'06)*, 740–744.
- Smaill, C. R. (2005). The Implementation and Evaluation of OASIS:A Web-Based Learning and Assessment Tool for Large Classes. *IEEE Transaction on Education*, 48 (4), 1–6.
- Taha, H. A. (1997). *Operations Research: An Introduction*, Upper Saddle River, NJ: Prentice-Hall.
- Takemura, Y., Shima, K., Matsumoto, K., Inoue, K., & Troii, K. (1999). Factor analysis of comprehension states in the learning phases of a programming language. *In Proceeding of the 1999 APSEC'99 International Conference on Software Engineering*, 136-143.
- Tsaganou, G., Grigoriadou, M., Cavoura, T., & Koutra, D. (2003). Evaluating an intelligent diagnosis system of historical text comprehension. *Expert Systems with Applications*, 25, 493-502.
- U. S. Department of Energy (2003). *Assessment Criteria and Guidelines for Determining the Adequacy of Software Used in the Safety Analysis and Design of Defense Nuclear Facilities*, retrieved July 15, 2008 from [http://www.hss.energy.gov/csa/csp/sqa/assessments/crads/sqa\\_CRAD\\_4.2.4.pdf](http://www.hss.energy.gov/csa/csp/sqa/assessments/crads/sqa_CRAD_4.2.4.pdf).
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*, New York: Springer-Verlag.
- Vapnik, V. (1998). The Support Vector Method of Function Estimation. In J. A. K. Suykens & J. Vandewalle (Eds.), *Non-linear Modeling: Advanced Black-Box Techniques*, Boston, MA: Kluwer, 55–85.
- Vrettaros J., Vouros G., Drigas A. S. (2004). Development of a Diagnostic System of Taxonomies Using Fuzzy Logic. *WSEAS Transactions on Information Science and Applications*, 1 (6), retrieved July 15, 2008 from [http://imm.demokritos.gr/publications/Development\\_Diagnostics.pdf](http://imm.demokritos.gr/publications/Development_Diagnostics.pdf).
- Wan, V., & Campbell, W. M. (2000). Support vector machines for speaker verification and identification. *In Proc. IEEE Workshop Neural Networks for Signal Processing*, Sydney, Australia, 775–784.
- Yang, T. (2000). Computational verb systems: a new paradigm for artificial intelligence. *Information Sciences*, 124 (1-4), 103-123.
- Yuan S. T., & Chen Y. C. (2008). Semantic Ideation Learning for Agent-Based e-Brainstorming. *IEEE Transactions on Knowledge and Data Engineering*, 20 (2), 261-275.
- Zhang, J., Cheung, B., & Hui, L. (2001). An intelligent tutoring system: SmartTutor. *Proceedings of 2001 World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2130-2131.
- Zhang, R., & Rudnick, A. I. (2002). Improve Latent Semantic Analysis Based Language Model by Integrating Multiple Level Knowledge. *In Proceeding of seventh International Conference on Spoken Language Processing*, 893-896.

## **Improving Recall and Transfer Skills Through Vocabulary Building in Web-Based Second Language Learning: An Examination by Item and Feedback Type**

**Seongchul Yun**

Department of Instructional System Technology, Korea National University of Education, Korea // Tel: +82-43-230-3431 // seongchul.yun@gmail.com

**Paul Chamness Miller**

Division of Teacher Education, University of Cincinnati, USA // Tel: +1 513-556-0794 // paul.miller@uc.edu

**Youngkyun Baek, Jaeyeob Jung**

Department of Instructional System Technology, Korea National University of Education, Korea // Tel: +82-43-230-3431 // youngkyun.baek@gmail.com // jaeyeob.jung@gmail.com

**Myunghwan Ko**

Kumsan High School, Korea // Tel: +82-41-753-9211 // blueidea@chol.com

### **ABSTRACT**

The purpose of this study is to investigate the effectiveness of response modes by item and feedback type in a web-based language learning program. The subjects of this study, 122 Korean tenth graders learning English as a foreign language, were placed into groups of four and were given a web-based language learning program consisting of two different varieties of item type (multiple choice and constructed response) and two forms of feedback (explicit and implicit). The results of this study suggest that the constructed response mode has a greater effect than the multiple-choice mode on the recall and the transfer of the students' learning. Second, the explicit response mode has a greater effect than the implicit response mode on the recall and the transfer of the students' learning. Third, there is an interactional effect between the item response mode and the feedback response mode in web-based language learning.

### **Keywords**

Web-based language learning, Recall, Transfer, Vocabulary acquisition

### **Introduction**

Research on learning through the web has largely focused on interactivity (Kahveci & Imamoglu, 2007; Stromso, Grottum, & Lycke, 2007; Wagner, 1997). It is commonly accepted that there are three types of interaction: interaction between the learner and the content being learned, interaction between the learner and the instructor, and interaction between learners (see Kahveci & Imamoglu, 2007, for a current review).

Interaction between the learner and the content presumes that the participation of the learner in the learning process serves as a reaction to the content. The development of web technology not only makes it possible for learners to recognize and react to the feedback presented, but such technology also makes various response modes possible (e.g., form in web-based documents). These various response modes serve as a critical medium for supporting the interaction between learners and content. Clariana & Lee (2001) point out that constructed responses are hard to manage technically and requires much effort due to the characteristics of the web. They argue that in web-based learning, constructed responses have decreased in usage and that multiple-choice formats have become more frequent; they further maintain that this trend will continue. However, along with this trend are some key points to keep in mind.

Response modes in the form of true-false, matching, and multiple-choice require recognition of related knowledge, while short answer, constructed, and completion modes require the recall of knowledge (Clariana & Lee, 2001).

Additionally, it is important to note that much of the previous research on this topic has revealed that recognition and recall are different memory retrieval processes and require different cognitive skills (Clariana & Lee, 2001; Godden & Baddeley, 1975, 1980; Martinez & Katz, 1996).

What is more, Glover's (1989) research suggests that recall is more effective in learning than recognition, and that multiple-choice in web-based learning tends to reinforce recognition, not recall. According to constructivist assumption stressing the active construction of knowledge and situational activities (see Han, 1990), recall is a much more important intellectual activity than recognition because recall makes retrieving schemata an easier process for the learner. This coincides with the context where students construct their knowledge through the representation of specific knowledge internally and the interpretation of personal experiences. In the same context, recall plays a greater role in transferring knowledge than recognition, because recognition is a process of simply retrieving what the learner has obtained previously or what the learner has memorized (Han, 1990).

Viewed from the fact that recall and recognition are two different cognitive processes and that recall is closely related to the transfer of knowledge, it is unfavorable to adopt a multiple response mode that requires recognition in web-based learning. Graff (2003) suggests that introducing design principles that incorporate the learners' cognitive style (i.e., holistic and analytic) with content might be instrumental in developing an effective instructional program. Thus, what is needed at this time is to test response modes that support active intellectual recall and the transfer of knowledge matching to the context. More specifically, it is the explicit response presented at the time when feedback is given that requires greater concern. We define the explicit response as an action offered after providing an answer to multiple response modes such as keyboard input, feedback, written answers from the students, among others.

In terms of response modes, in his summary of several related studies, Tobias (1973) concluded that the constructed response is the most effective for all academic achievements. He also found that the explicit response results in better achievement than the implicit response, allowing more time for students to reflect and engage in a short reading after responding to multiple-choice items. According to Clariana & Lee (2001), the multiple-choice response, which they presented as an explicit response, received a higher achievement score than the constructed response in learning vocabulary in the field of instructional design. On the effectiveness of various response modes, familiarity and difficulty are the variables that had the greatest impact on learning. In terms of unfamiliar or difficult content, the constructed and the explicit response modes were more effective than the multiple-choice and implicit response modes on achievement, while there was no significant difference between response modes when students had a higher level of familiarity with the content (Clariana & Lee, 2001; Tobias, 1973).

Another importance in the design of language learning lies in the level of information processing at the lexical level. Koda (1996) asserts that in the process of reading, the reader constructs meaning from the information provided in the text. She further posits that reading comprehension occurs as a result of the interaction between the reader's own pre-existing knowledge (i.e., schemata) and the information provided in the text. In terms of vocabulary learning, it is believed that the rate of memorized words is predicted to be determined by the level of information processing (Liu, 1992), and is thus connected to the process of reading (Koda, 1996). This implies that vocabulary acquisition should be organized in a meaningful way rather than just in repetition.

Through reviews of the literature on recall and recognition in web-based learning, one might deduce that the constructed response mode is more effective than the multiple-choice mode, the explicit response mode has an effect on the recall and the transfer of the knowledge, and the type of response mode makes a difference in effectiveness according to the difficulty level and familiarity of the learning tasks. The present study contributes to the theoretical background of the effectiveness of response modes, considering cognitive traits of learners and levels of learning tasks by testing the effects of various response modes on recall and transfer of learning according to the level of the learning task's difficulty and information processing.

### **Response Modes in Web-Based Learning**

Generally, the term *response* in web-based language learning is understood as the feedback provided as a result of learners' responses to items in the activity. In this instance, the computer keeps the result of responses (i.e., an answer or a selection among the examples provided), and such responses are either an action of the student or an indication for one among provided examples. Gredler (1999) defined response modes as a way of answering a

question or a reaction to the learning materials presented to them. She suggests that there are two types of response modes: a multiple-choice response, where the learner merely *selects* the answer, or a constructed response, where the student *produces* an answer. In a web-based environment, a multiple-choice response is a type of selective frame providing the student with numerous examples, from which the student chooses the right answer. A constructed response would require the learner to construct an answer, such as a sentence, paragraph, essay, etc., calculating a mathematical problem or drawing a diagram.

## Response Modes by Item Type

Usually items for questioning are grouped into either a selection type or a constructed type (Gredler, 1999). The selection type includes true-false, matching, and multiple-choice items. This type of questioning is sometimes referred to as *recognition* because it requires the recognition process of confirming one selection among several possible answers (Grunlund, 1993). True-false items give a statement to the students, requiring that they judge whether the statement is true or false. Web-based true-false items are typically made of a radio button form field. Matching the type of item is a questioning method which provides both a series of presuppositions and answers, asking students to make an appropriate match between them, such as terminology, sentence, and symbols which could be used for either the presupposition or the answer. Web-based matching items are mainly composed of a check box form field and multiple-choice items are made up of an item stem and two or more answers from which to select. The item stem consists of an interrogative or incomplete sentence. On the web, when there is one correct answer, a radio button is used. When there are two correct answers such as the multiple-response variety, multiple-choice responses with check box form fields are adopted.

Table 1. Example of a multiple-choice item

Example of a multiple- choice item	<p>Which of the following is a chemical transition?</p> <p>① Water frozen to form ice.</p> <p>② A piece of paper burnt into ash.</p> <p>③ A rock smashed into pieces.</p> <p>④ Salt dissolved in water resulting in salt water.</p>
Example of a multiple- choice item on the web	<p>Which of the following is a chemical transition?</p> <p><input type="radio"/> Water frozen to form ice.</p> <p><input type="radio"/> A piece of paper burnt into ash.</p> <p><input type="radio"/> A rock smashed into pieces.</p> <p><input type="radio"/> Salt dissolved in water resulting in salt water.</p>

A selection type of item is characterized by its structuredness and objectification. Its main objective is to determine the student's ability to identify, select, and recognize the correct answer among several choices. In this case, the instructional designer makes up items inclined to target memorization and fragmentary knowledge.

A constructed type of item includes short answer, completion and essay formats, and is often referred to as the subjective type of response. This is characterized by expecting the student to type in the correct answer. Such items are also referred to as recall in that they require the learner's retrieval process to generate an answer without any prompts (Grunlund, 1993). Short answer items require such answers as a short sentence, a phrase or numbers. The web-based version of this is made up of a one-line text box with a form field (see Table 2 below).

The completion type of item is similar to a cloze activity that contains a blank in a sentence, figure, or diagram to be completed by the student. The web-based type of completion is composed of a one-line text box with a form field.

Table 2. Example of a short answer

Example of a short answer type	1. Who wrote Hamlet? 2. List three types of reliability.
Example of a short answer type on the web	1. Who wrote Hamlet? <input type="text"/> 2. List three types of reliability. <input type="text"/> <input type="text"/> <input type="text"/>

Table 3. Example of a web-based completion item

Example of a completion type	1. The author who wrote Hamlet is _____.
Example of a completion type on the web	1. The author who wrote Hamlet is <input type="text"/>

An essay item requires the student to write a minimum of one paragraph, given a particular situation or topic. The web-based version is made up of a scrolling text box with a form field.

Table 4. Example of a web-based essay item

Example of essay type	In 100 words or less, write about the necessity of computers in our daily life.
Example of essay type on the web	In 100 words or less, write about the necessity of computers in our daily life. <input type="text"/>

Although the retrieval process of information is different, depending on the item type, the multiple-response mode rather than the constructed response mode is more frequently adopted in instructional situations. Clariana & Lee (2001) argue that the use of the constructed response mode has decreased in use compared to the multiple-choice response mode and this trend will continue in web-based learning. They maintain that the reason for the decreased use of the constructed response mode is due to the technical difficulties of creating this type of item on the web. In particular, it is difficult to score constructed responses because they cannot be automatically identified as correct or incorrect like multiple-response items. In contrast, the multiple-choice response mode, which is a type of recognition test, is preferred in web-based learning due to its easy implementation and fast loading on the web. However, there are two main reasons why the constructed response mode, a type of recall test, should be implemented in web-based learning. The first reason is that recalling and recognizing produce different learning outcomes of declarative knowledge (Clariana, 2003). The second reason is that much of the research supports the notion that recalling tasks are more effective learning tools than recognition tasks (Clariana & Lee, 2001; Tobias, 1973).

### Response Modes by Feedback Type

One of the advantages the computer has for instructional designers is that it supports the elaboration of stimuli in digital formats and their presentation to students. Another advantage is that the computer elicits a response from the student and provides immediate, detailed and relevant feedback (Gagné & Wagner, 1988). Most feedback contains information on the adequacy or inadequacy of the students' performance. As in the classroom, web-based learning can also provide various types of feedback (i.e., implicit and explicit). However, the difference is that the web-based context may reduce the affective filter (see Krashen, 1982) that the learner may have in the potentially coercive environment of teacher driven instruction.

Richards (1989) tested the effects of explicit and implicit feedback provided by the computer. In this study, the explicit response is the constructed response, where following the erroneous response the student is prompted to retype the correct answer in response to the feedback pertaining to the explanation of the correct response. The implicit response provides feedback without the opportunity for the learner to provide the correct response. The

results indicate that explicit feedback followed by the opportunity for the learner to provide the correct answer results in higher academic achievement than implicit feedback followed by the provision of the correct answer in the recall and recognition post-tests. Therefore, one might conclude that by leading the learner to the correct answer while providing explicit feedback may increase the academic achievement of students in web-based learning. This result is similar to the findings of Tobias (1973) and Clariana & Lee (2001).

## Research Questions

As the above review of the literature has indicated, there are several questions that have arisen in previous work examining the effectiveness of response modes and feedback types on Korean students' recall and transfer of English vocabulary. This study focuses on the following research questions: (1) Are there any differences in the recall and transfer of vocabulary between item types in a web-based language learning program? (2) Are there any differences in the recall and transfer between explicit and implicit feedback types in a web-based language learning program? (3) Are there any interactional effects between item types and feedback types on the recall and the transfer of English vocabulary?

## Research Methods

### Research Subjects

One hundred twenty two tenth graders, with an average age of 16 years, from a large metropolitan high school in South Korea were the subjects of this study. The participants came from 4 separate classes in the same school. The participants were randomly assigned into two groups for the various tasks in the study.

### Recall Pre-test

The pre-test for the recall of English words was administered so as to determine the homogeneity of the groups. The pre-test consisted of 20 items from the Korean National Standard English Curriculum and the maximum score was 20 points. These tests are modeled after Han (1990) and Yoon (1997)'s test around vocabulary in a Web-based language learning program.

The results of the pre-test are presented in Tables 5 & 6. These tables show that groups are homogeneous in terms of the recall test.

*Table 5. t-test for Recall Pre-test Means of the Two Item Types*

Group	N	Mean	SD	df	<i>t</i>	P
Constructed	60	9.03	4.770	120	720	473
Multiple-choice	62	9.70	5.466			

*Table 6. t-test for Recall Pre-test Means of the Two Feedback Types*

Group	N	Mean	SD	df	<i>t</i>	P
Explicit	59	9.44	5.396	120	1.486	140
Implicit	63	8.06	4.836			

### Instructional Events and Flow in the Web-based Language Learning Program

A Web-based language learning program was developed with the specifications described below. The design principles and flow of the web-based language learning program were based on Gagné's (1985) instructional events. The program was developed using the following design principles, forms, and media formats (see Figures 1 and 2).

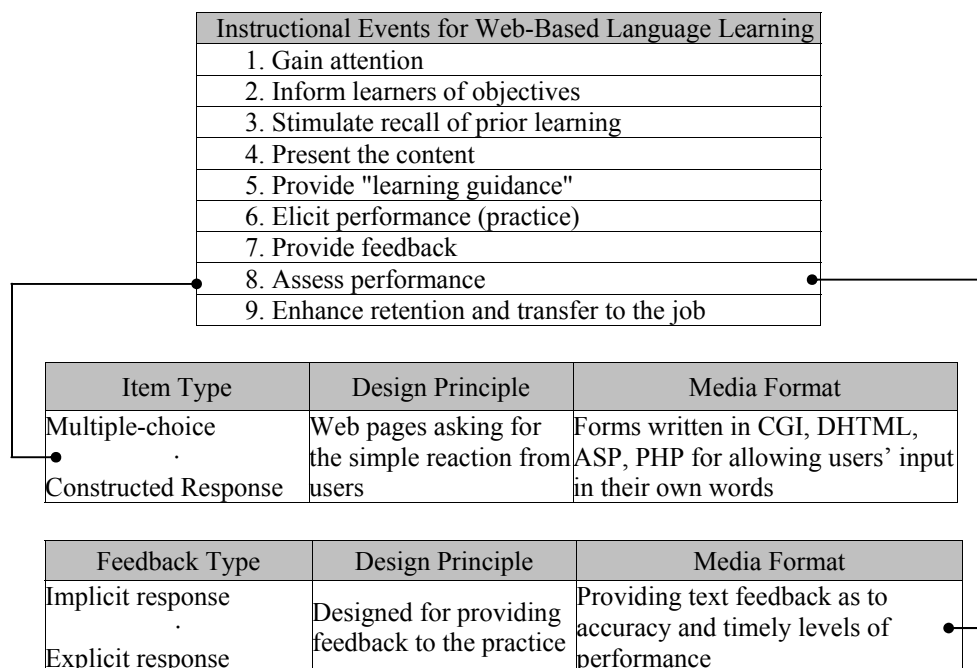


Figure 1. Principles and Media for Response Modes

The flow chart of the Web-based language learning program is presented in Figure 2.

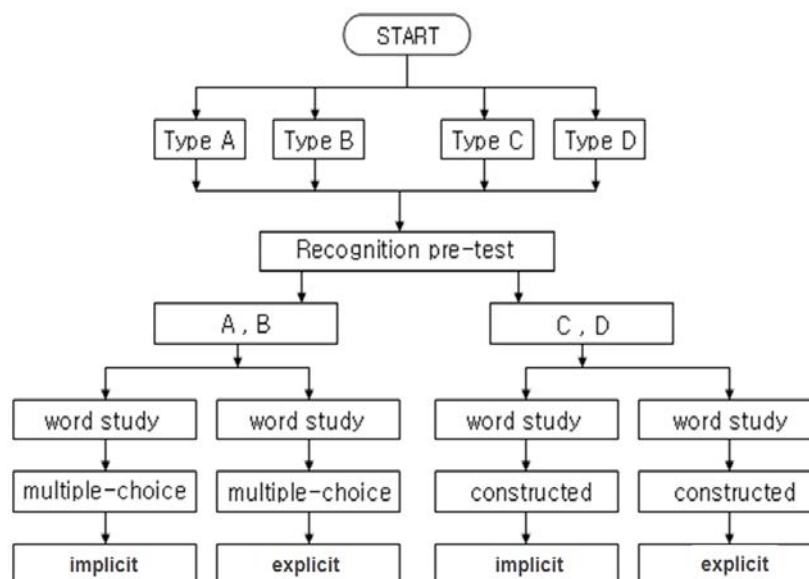


Figure 2. Flow of Learning in Web-Based Language Learning

### Learning Material and Sample Screens in the Web-based Language Learning Program

The content for this study consisted of 20 items from the Korean National Standard English Curriculum for tenth graders. The words were presented in two formats, ten each in a dictionary format and in a contextual format. The dictionary format was adopted from Naver's English Dictionary (<http://endic.naver.com>) and the contextual format was adopted from Yoon (1997) which includes the word in a sentence.

The beginning screens for the web-based program are presented in Figure 3. When the user logs in, the screen presents 4 tests from which to choose.



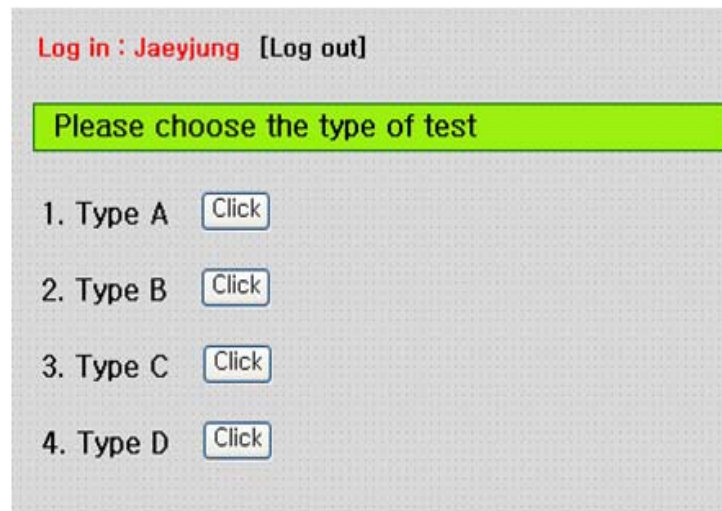


Figure 3. Log-in and Options Screen

Clicking one option on the screen leads the user to the recall test consisting of 20 items as shown in Figure 4. To avoid potential user problems and unnecessary actions, Internet Explorer's tool box and the use of the right mouse button are blocked. If the user does not answer a question and tries to skip an item, a warning pop-up window opens, prompting the user to answer the item.

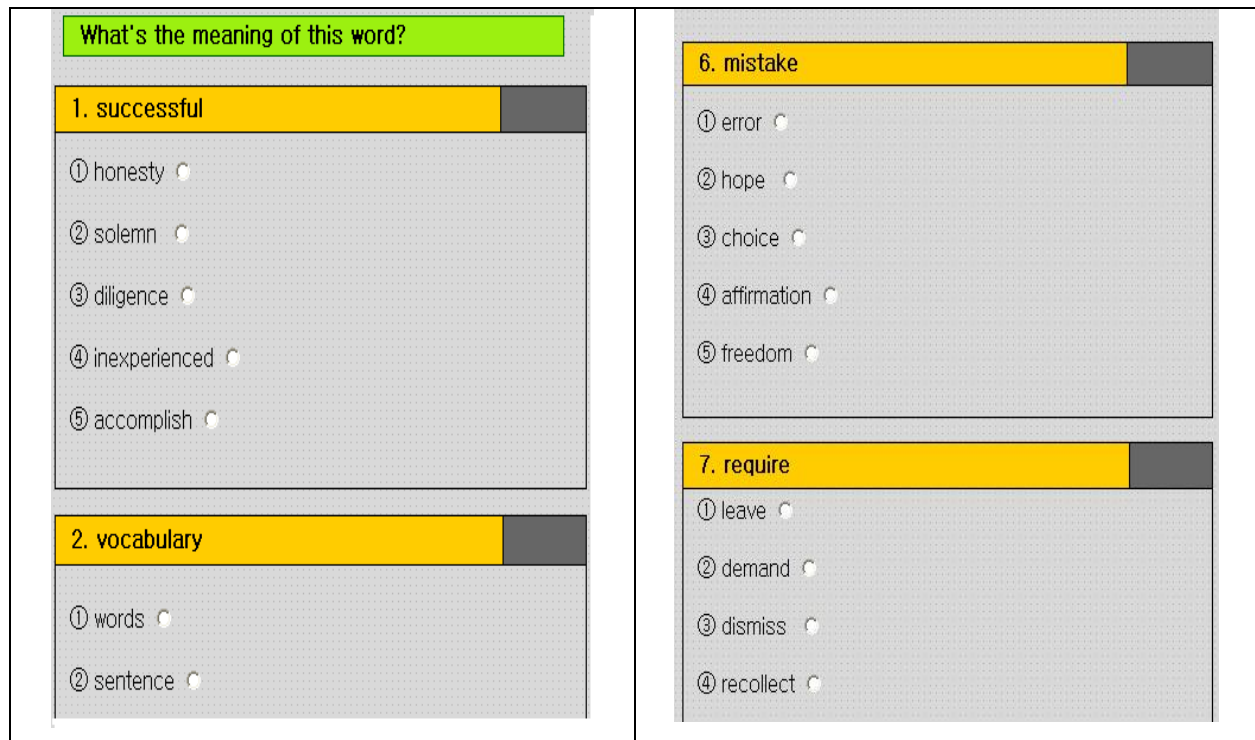


Figure 4. Screens for Recall Test

After the recall test takes place, the instructional component of the program begins. Ten words in the dictionary format followed by 10 words in the contextual format are presented as shown in Figure 5.

The figure shows two screenshots of a software interface. The top screenshot, titled 'Word study', displays the word 'successful' as an adjective. It provides a definition: '1 (successful) having succeeded or being marked by a favorable outcome', followed by two bullet points: '• a successful architect' and '• a successful business venture'. A 'next' button is at the bottom. The bottom screenshot, titled '11.', shows three sentences: 'My wife decided I should have my *own* shop.', 'He could no longer trust his *own* judgement.', and 'His office had its *own* private entrance.' Below the sentences, the word 'own' is defined as '(own, have, possess) have ownership or possession of'. A 'next' button is at the bottom.

Figure 5. Screens for the Dictionary Format (upper) and the Contextual Format (lower)

Next, the user answers a series of questions through input forms on the screen according to the response modes as shown in Figure 6. In order to control the response time for all participants and reduce the potential threat of validity, each participant has a time limit of 30 seconds to respond to each item. After 30 seconds, the program proceeds to the next screen regardless of the user's response.

The figure shows two side-by-side screenshots of a software interface. The left screenshot, titled 'What's the meaning of this word?', shows the word '1. successful' with a 25-second timer. Below it are five multiple-choice options, each with a radio button: ① honesty, ② solemn, ③ diligence, ④ inexperienced, and ⑤ accomplish. A 'next' button is at the bottom. The right screenshot, also titled 'What's the meaning of this word?', shows the word '1. successful' with a 25-second timer. Below it is a large empty text box for a constructed response. A 'next' button is at the bottom.

Figure 6. Screens for Multiple-Choice (left) and Constructed Responses (right)

During the feedback stage, either the explicit response mode or the implicit response mode is presented according to the selected option. In the explicit response mode, the full explanation is presented again (see Figure 7).

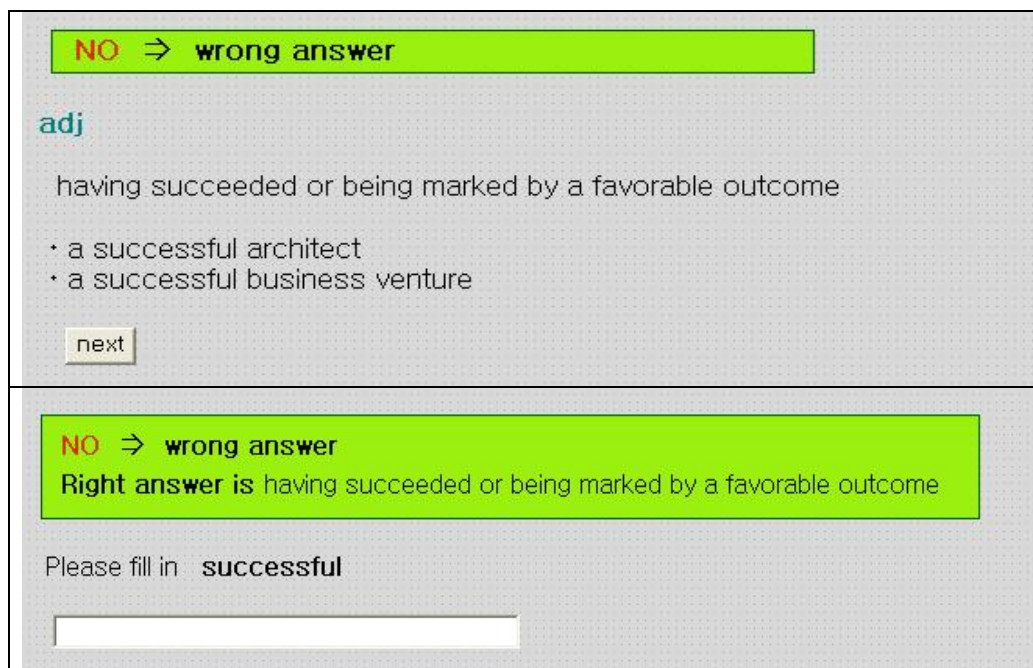


Figure 7. Screens for the implicit response mode (upper) and the explicit response mode (lower)

### Post-tests for Recall and Transfer

When the instructional session ended, the recall and transfer post-tests were administered. These tests were modeled after Han (1990) and Yoon (1997)'s test around vocabulary in a Web-based language learning program. The recall post-test was a paper-based test where the user was asked to write a definition of the presented vocabulary. The transfer post-test was also paper-based, and asked the user to write new sentences using the vocabulary presented during the instructional session. The students were required to write their selected words in sentences keeping the appropriate context in mind. The items used for the post-test were 20 vocabulary words, with a maximum score of 20. In order to measure the internal consistency reliability of the post-test, the Cronbach' Alpha was applied, with .91 for the recall post-test, and .86 for the transfer post-test. These results clearly exceed the standard level of "acceptability" for this type of test (i.e., .70).

## Results and Interpretation

### Descriptive Analysis

The participants were randomly assigned to the groups receiving the various treatments. Tables 5 and 6 show the means and standard deviations of each group for the recall and transfer tests.

Table 5. Means and Standard Deviations of the Recall Test by Group

Group	Multiple	Constructed	Implicit	Explicit
N	62	60	63	59
Mean	9.60	12.28	9.05	12.92
Standard Deviation	5.880	5.396	5.144	5.799

Table 5 shows that the explicit group has a mean of 12.92 which is the highest, whereas the implicit group has a mean of 9.05 which is the lowest in the recall test. This result implies that the explicit response is more effective than the implicit response in the recall test.

Table 6. Means and Standard Deviations of the Transfer Test by Group

Groups	Multiple	Constructed	Implicit	Explicit
N	62	60	63	59
Mean	5.60	8.03	5.38	8.31
Standard Deviation	4.779	4.547	4.419	4.775

Table 6 shows that the explicit group has a mean of 8.31 which is the highest, whereas the implicit group has a mean of 5.38 which is the lowest in the transfer test.

### Differences in Recall and Transfer Between Item Types

The difference in the recall test between the item types is presented in Table 7. The *t*-test procedure was applied for the mean scores of the constructed and the multiple response groups.

Table 7. *t*-test for Recall Test Means of the Two Item Types

Group	N	Mean	SD	df	<i>t</i>	p
Constructed	60	12.28	5.40	120	2.627	.01
Multiple-choice	62	9.60	5.88			

Table 7 shows that there is a significant difference in the recall test score between the two response modes in the web-based language learning program, which suggests that the constructed response mode has a greater effect than the multiple-response mode on the recall in the web-based language learning program.

The difference in the transfer test between the item types is presented in Table 8. A *t*-test procedure was applied for the mean scores of the constructed and the multiple response groups.

Table 8. *t*-test for Transfer Test Means of the Two Item Types

Group	N	Mean	SD	df	<i>t</i>	p
Constructed	60	8.03	4.55	120	2.883	.005
Multiple-choice	62	5.60	4.78			

As table 8 shows, there is a significant difference in the transfer test score between the two response modes in the web-based language learning program. This result indicates that the constructed response mode has a greater effect than the multiple-response mode on transfer in the web-based language learning program.

### Differences in the Recall and Transfer Between Feedback Types

The difference in the recall test between the feedback types is presented in Table 9. A *t*-test procedure was applied for the mean scores of the explicit and implicit response groups.

Table 9. *t*-test for Recall Test Means of the Two Feedback Types

Group	N	Mean	SD	df	<i>t</i>	p
Explicit	59	12.92	5.80	120	3.902	.000
Implicit	63	9.05	5.14			

It is clear from the data presented in table 9 that there is a significant difference in the recall test score between the two feedback types in the web-based language learning program. The results suggest that explicit feedback has a greater effect than implicit feedback on recall in the web-based language learning program.

The difference in the transfer test between the feedback types item is presented in Table 10. A *t*-test procedure was applied to the mean scores of the explicit and implicit response groups.

Table 10. *t*-test for Transfer Test Means of the Two Feedback Types

Group	N	Mean	SD	df	<i>t</i>	p
Explicit	59	8.31	4.78	120	3.513	.001
Implicit	63	5.38	4.42			

As is indicated in table 10, there is a significant difference in the transfer test score between the two feedback types in the web-based language learning program, which suggests that explicit feedback has a greater effect than implicit feedback on transfer in the web-based language learning program.

### Interactional Effect between Item Types and Feedback Types on Recall and Transfer

The interactional effect between the item types and the feedback types on the recall test score in the web-based language learning program was tested using a two-way ANOVA. The results of the analysis are shown in Table 11.

Table 11. Two-Way ANOVA for Recall Scores of the Item and Feedback Types

Source of Variance	SS	df	MS	F	p
Item types	210.935	1	210.935	7.658	.007
Feedback types	443.019	1	443.019	16.083	.000
Item x Feedback	140.796	1	140.796	5.111	.026
Error	3250.317	118	27.545		
Total	4047.180	121			

Table 11 shows that there is an interactional effect between the item types and the feedback types on the recall test score in the web-based language learning program. As shown in Figure 8, the constructed response combined with the explicit response has the highest recall score, whereas the multiple-choice response has the lowest recall score.

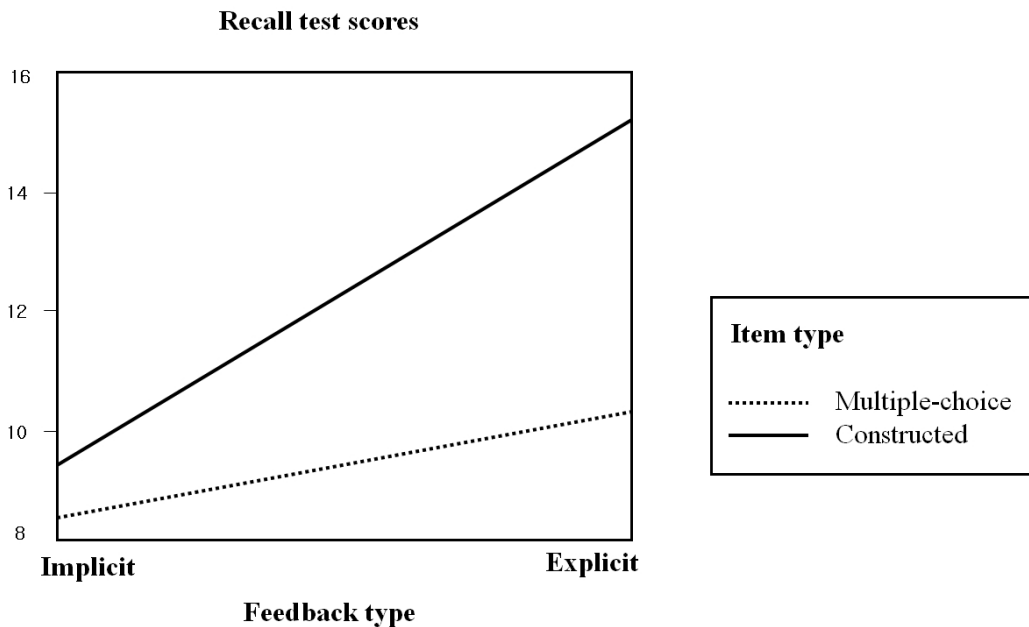


Figure 8. Interaction Effect of the Item Type and Feedback Type on Recall

To test the interactional effect between the item types and the feedback types on the transfer test score in the web-based language learning program, a two-way ANOVA procedure was applied. The results of the analysis are shown in Table 12.

Table 12. Two-way ANOVA for Transfer Scores of the Item and Feedback Types

Source of Variance	SS	df	MS	F	p
Item types	177.662	1	177.662	9.480	.003
Feedback types	252.510	1	252.510	13.473	.000
Item x Feedback	154.413	1	154.413	8.239	.005
Error	2211.498	118	18.742		
Total	2793.877	121			

Table 12 shows that there is an interactional effect between the item and feedback types on the transfer test score in the web-based language learning program. As shown in Figure 9, the constructed response combined with the explicit response has the highest transfer score, whereas the multiple-choice response has the lowest transfer score.

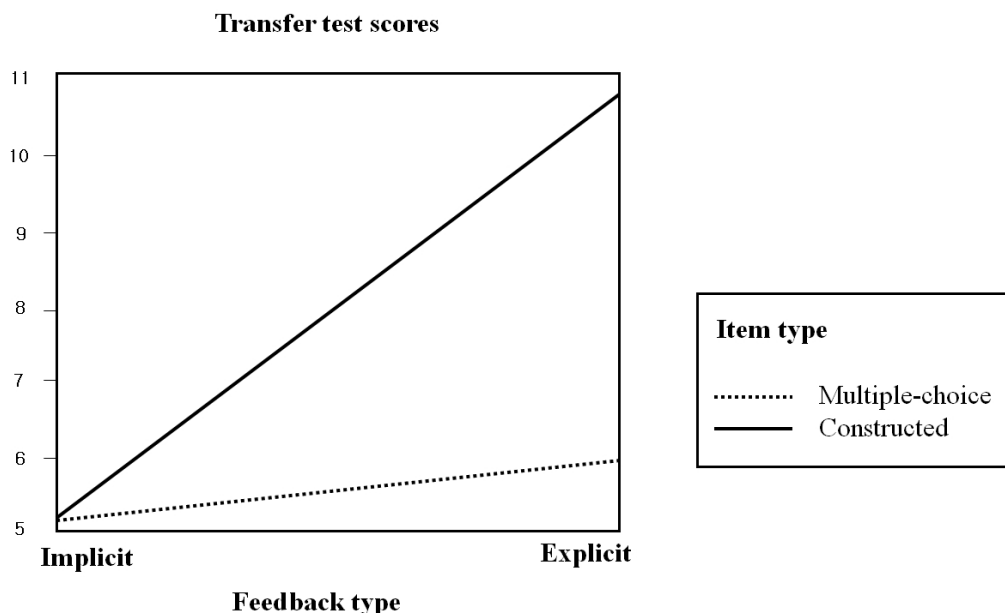


Figure 9. Interaction Effect of the Item & Feedback Types on Transfer

## Discussion

This study sought out to investigate the effectiveness of response modes by the type of item difficulty and information processing levels in a web-based English language learning program. Also this study focused on analyzing the interaction between the item and feedback types which effect students' performance. Based on the results of the data presented above, the following are the implications for this study.

First, there is a difference between groups in the effect of the response modes on the students' recall. Most importantly, the constructed-explicit response has the biggest effect on the students' recall in the web-based language learning program. In addition there is a difference in the effect of the response modes on the students' transfer in the web-based language learning program. The constructed-explicit response has the largest effect on the students' transfer. These findings support the findings of other research in vocabulary acquisition. For example, Nation (1990) argues that both direct instruction (such as the explicit feedback provided in this web-based program and contextual learning (such as the constructed response mode used in this web-based program) are useful in promoting vocabulary learning. This claim is further supported by Paribakht & Wesche (1997), who maintain that a combination of contextual learning and direct instruction is a better approach to vocabulary acquisition than context alone.

Second, there is a difference in the effect of the item types on the students' learning. That is to say, the constructed response mode has a higher effect on the learners' recall and transfer than the multiple-choice response mode. This supports the notion that the multiple-choice response mode facilitates recognition and that the constructed response

mode facilitates recall. These findings also suggest that a deeper examination of the theory of vocabulary acquisition is needed in order to fully understand such concepts as “receptive” and “productive” vocabulary knowledge (see Coady, 1993; Grabe & Stoller, 1997).

Third, there is a difference in the effects of the feedback types on students’ learning. More specifically, the explicit response mode has a greater effect on recall and transfer than the implicit response mode. These findings suggest that the explicit response mode draws an active response from the students, and thus helps to activate their knowledge. The issue of feedback in the language learning context is a topic that is also highly debated, with empirical evidence to both support and refute its effectiveness (see Norris & Ortega, 2000, for a meta-analytic review of feedback). The bulk of the research on feedback is on oral corrective feedback (typically in teacher/student and native/non-native speaker settings), whereas the present study contributes to the discussion by focusing on computer-generated feedback. For a review of the significant research on corrective feedback in the language learning context, see Miller (in press).

Fourth, there is a significant interaction between the item types and the feedback types on students’ recall and transfer, with the constructed-explicit response mode having the greatest effect. This result indicates that the response mode which requires an active response from the learner affects the recall and the transfer of the students’ learning. This finding also supports the ideas of contextualized learning and direct instruction found in the body of research (see Nation, 1990 and Paribakht & Wesche, 1997).

## **Conclusion**

As was noted in the results and further discussed above, several conclusions were reached in analyzing the data collected. From this analysis, we were able to address the research questions initially posed at the onset of the study.

### **Are there any differences in the recall and transfer of vocabulary between item types in a web-based language learning program?**

In considering the first research question, the results suggest that constructed responses have a greater effect than multiple-choice responses on the recall and transfer of students in a web-based language learning environment. This conclusion supports the information processing theory on recall and transfer.

### **Are there any differences in the recall and transfer between explicit and implicit feedback types in a web-based language learning program?**

The explicit response mode as feedback has a greater effect than the implicit response mode on the recall and transfer of vocabulary in a web-based language learning environment. This implies that response modes requiring the active participation of the students can affect academic achievement more than just clicking along a given path in a web-based language learning environment.

### **Are there any interactional effects between item types and feedback types on the recall and the transfer of English vocabulary?**

Based on the results of this study, one can conclude that there is an interactional effect between the item types and the feedback types on the recall and transfer of vocabulary in a web-based language learning environment. It is especially interesting to note that the constructed-explicit response mode maximizes the effects on recall and transfer.

## **Limitations & Future Research**

Although this study clearly contributes to our understanding of recall and transfer in web-based vocabulary learning, there are limitations to consider. First, the focus of the study was on testing for statistical significance. Future research should consider mixed designs or studies that examine qualitative aspects of the topic. In terms of research design, there is also a limitation in that the participants were forced to provide a response, even if that meant

guessing at the right answer. The potential error that this produced in the results could not be treated through statistical analysis.

Additionally, in order to understand the reading process better (including vocabulary acquisition), there are several questions that future research should consider. The first is to explore the effect of different learning tasks in a web-based language learning program, considering the specific characteristics of the tasks. This study focused only on the item difficulty level. Although the focus of this study was on the memorization of vocabulary, one might compare the results between different types of vocabulary, such as content and function words and receptive and productive vocabulary (Aebersold & Field, 1997). The frequency of the vocabulary, or whether it is “core” to the genre being learned, is another area of future research (Grabe & Stoller, 1997). It is also important to consider how web-based language learning programs might impact vocabulary acquisition in the reading context (e.g., comparing the effect it may have in the top-down and bottom-up approaches to reading). Future research should also consider the effect web-based vocabulary learning may have in the context of direct versus indirect instruction (see Nation & Newton, 1997). Other types of response mode may also be considered in future studies on web-based language learning.

## References

- Aebersold, J. A. & Field, M. L. (1997). *From reader to reading teacher: Issues and strategies for second language classrooms*, Cambridge: Cambridge University Press.
- Clariana, R. B. (2003). The effectiveness of constructed-response and multiple-choice study tasks in computer aided learning. *Journal of Educational Computing Research*, 28, 395-406.
- Clariana, R. B., & Lee, D. (2001). The Effects of recognition and recall study tasks with feedback in a computer-based vocabulary lesson. *Educational Technology Research and Development*, 49 (3), 23-36.
- Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In Huckin, T., Haynes, M. & Coady, J. (Eds.), *Second language reading and vocabulary learning*, Norwood, NJ: Ablex, 3-23.
- Gagné, R. M. (1985). *The condition of learning*, New York: Holt, Rinehart and Winston.
- Gagné, R. M., & Wagner, W. (1988). *Principles of instructional design*, New York: Rinehart & Winston.
- Glover, J. A. (1989). The testing phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, 81, 392-399.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66, 325-331.
- Godden, D. R., & Baddeley, A. D. (1980). When does context influence recognition memory? *British Journal of Psychology*, 71, 99-104.
- Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language. In Coady, J. & Huckin, T. (Eds.), *Second language vocabulary*, Cambridge: Cambridge University Press, 98-122.
- Graff, M. (2003). Learning from web-based instructional systems and cognitive style. *British Journal of Educational Technology*, 34, 407-418.
- Gredler, M. E. (1999). *Classroom assessment and learning*, New York: Longman.
- Grunlund, N. (1993). *How to make achievement tests and assessments* (5<sup>th</sup> Ed.), New York: Allyn and Bacon.
- Han, H. (1990). *A study on the effects of sentential context learning on recall and recognition of English vocabulary*, Unpublished master's thesis, Hanyang University, Seoul, South Korea.
- Kahveci, M., & Imamoglu, Y. (2007). Interactive learning in mathematics education: Review of recent literature. *Journal of Computers in Mathematics and Science Teaching*, 26 (2), 137-153.
- Koda, K. (1996). L2 word recognition research: A critical review. *The Modern Language Journal*, 80, 450-460.
- Krashen, S. (1982). *Principles and practice in second language acquisition*, New York: Pergamon.



- Liu, M. (1992). Hypermedia-assisted instruction and second language learning: A semantic-network-based approach. *Paper presented at the Annual Conference of the Eastern Educational Research Association*, Hilton Head, SC, (ERIC Document Reproduction Service No. ED349954).
- Martinez, M. E., & Katz, I. R. (1996). Cognitive processing requirements of constructed figural response and multiple-choice item in architecture assessment. *Educational Assessment*, 3, 83-98.
- Miller, P. C. (in press). Corrective feedback in the L2 classroom: Where are we now? In Mantero, M., Miller, P. C. & Watzke, J. (Eds.), *Readings in language studies: Language across disciplinary boundaries* (Vol. 1), New York: International Society for Language Studies.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*, New York: Newbury House.
- Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In Coady, J. & Huckin, T. (Eds.), *Second language vocabulary acquisition*, Cambridge: Cambridge University Press, 238-254.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and qualitative meta-analysis. *Language Learning*, 55, 417-529.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In Coady, J. & Huckin, T. (Eds.), *Second language vocabulary acquisition*, Cambridge: Cambridge University Press, 174-200.
- Richards, D. R. (1989). A comparison of three computer-generated feedback strategies. *Eleventh Annual Proceedings of Selected Research Presentations of the Association for Educational Communications and Technology*, 11, 358-367 (ERIC Document Reproduction Service No. ED308831).
- Stromso, H. I., Grottum, P., & Lycke, K. H. (2007). Content and processes in problem-based learning: A comparison of computer-mediated and face-to-face communication. *Journal of Computer Assisted Learning*, 23, 271-282.
- Tobias, S. (1973). Review of the response mode issue. *Review of Educational Research*, 43, 193-204.
- Wagner, E. D. (1997). Interactivity: From agents to outcomes. In Cyrs, T. E. (Ed.), *New directions for teaching and learning* (Vol. 71), San Francisco: Jossey-Bass, 19-32.
- Yoon, M. (1997). *A comparative study on the effects of two vocabulary teaching methods in the EFL situation: The definitional approach and the contextual approach*, Unpublished Master's thesis, Seoul National University, Seoul, South Korea.

## A Computer System of Referential Resolution to Assess Students' Reading Comprehension

**Yu-Fen Yang**

Graduate school of applied foreign languages, National Yunlin University of Science & Technology, Douliu, Yunlin, Taiwan // Tel:+ 886-5-534-2601 Ext. 3136/Fax: +886-5-531-2058// yangy@yuntech.edu.tw

**Wing-Kwong Wong**

Graduate school of computer science & information engineering, National Yunlin University of Science & Technology, Douliu, Yunlin, Taiwan // Tel:+ 886-5-534-2601 Ext. 4330/Fax: +886-5-531-2063// wongwk@yuntech.edu.tw

**Hui-Chin Yeh**

Graduate school of applied foreign languages, National Yunlin University of Science & Technology, Douliu, Yunlin, Taiwan // Tel:+ 886-5-534-2601 Ext. 3106/Fax: +886-5-531-2036// hyeh@yuntech.edu.tw

### ABSTRACT

The major purpose of this study is to assist college students in constructing mental representation of referential resolution in a computer system. Ninety students were asked to draw mental maps indicating the relationships between references, take reading comprehension tests, and fill out open-ended questionnaire. Results showed that the correlation between referential resolution and reading comprehension ranged from 61% to 75% in three reading tasks. The more students practiced in referential resolution, the higher average score they obtained in the reading comprehension test. Students also requested fewer feedbacks when they progressed to better comprehend the final online text. That is, they became more independent and relied less on the system's feedback in their reading development. Although seventy-six students expressed that they learned from building connections of related words and made progressed in reading comprehension, fourteen students revealed that they still had difficulties after using the online system. This deserved more investigation in the future.

### Keywords

Referential resolution, Reading comprehension, Referring strategy, Mental representation, and Metacognition

### Introduction

College students who learn English as a Foreign Language (EFL) in Taiwan were found to experience great difficulty in resolving references (Chen, 2001; Yu, 1993; Chen & Dai, 2003). Results of Bensoussan and Laufer's study (1984) indicated that the major reading difficulty that ESL or EFL college students encountered was their failure to recognize the interrelationship between sentences in a text. In addition, Chu, Swaffar, and Charney (2002) mentioned that most Taiwanese EFL students were found to be less aware of how to use cohesive devices to integrate textual information. Wang and Ding (1998) further claimed that Taiwanese students' lack of skills in cohesive ties was sure to hinder them from comprehending texts.

Five cohesive ties were proposed by Halliday and Hasan (1976) and these ties were considered the keys to help readers integrate the meanings of the sentences in a text; they were reference, substitution, ellipsis, conjunction, and lexical cohesion. Among these five cohesive ties, reference accounted for 59% of variance in L1 readers' comprehension (Demel, 1990) and 75% of EFL students' statistical text comprehension (Huang, 2005). Three types of references, personal, demonstrative, and locative pronouns were considered in this study since they appeared more frequently in texts (e.g., Fortane, 2004; Kennison, 2003). Personal references refer to individuals or objects by specifying their functions or roles in the speech situation (Halliday & Hasan, 1976), such as "I", "me", and "you." Demonstrative references substitute nouns when the nouns can be understood from the context. They also indicate whether they are replacing singular or plural words. Examples include "this," "these," "that," "more," "neither," etc. Locative references are used to indicate locations. Examples include "here" and "there."

Once a reader received trainings in cohesive ties, he was supposed to identify the critical links in integrating successive sentences and forming a coherent discourse representation (Tea & Lee, 2004; Potelle & Rouet, 2003). This representation illustrated the reader's cognitive process of understanding a text. The reader might later use the

mental representation to summarize, analyze, discuss, and evaluate the information in a text (Guerrero, 2003). In Sharp's review (2003), different kinds of mental maps were built by the reader to organize knowledge and represent his/her cognitive structure, such as semantic networks, roundhouse diagrams, compare/contrast matrix, continuum/scale, and concept maps. According to Gardner (2004), mental maps provided students with alternative paths leading to the comprehension of a text. Reading the linear structure of a text was no longer the only way to comprehend a specific content. Instead, students were allowed to understand the content through the radial structure of diagrams represented in their mental maps. These mental maps not only displayed specific relationships among ideas but also enabled the reader to generate elaborative inferences about the text (Cuevas, Fiore, Bowers, & Salas, 2004).

The mental map could be viewed as the mental product of text comprehension for it helped the reader remove lexical or syntactic ambiguities, determine anaphoric antecedents or improve subsequent recall of information in long texts (Payne & Reader, 2006). In presenting students' mental maps, Greene and Azevedo (2007) suggested that self-regulated learning (SRL) processes accounted for qualitative shifts in students' mental maps from pretest to posttest in hypermedia learning environments (HLEs). In their study, the results indicated that students who exhibited a qualitative shift employed different processes, including metacognitive monitoring activities and learning strategies. They further claimed that "future research should focus on the best means of inculcating effective SRL behaviors through on-line methods, so that HLEs can teach both content and the actual process of learning" (p. 141). This led to substantive recommendation for the use of trace logs and scaffolding to promote self-regulated learning.

In Schacter, Herl, Chung, Dennis, and O'Neil's study (1999), four computational tools were designed to support assessment of students' problem-solving performance: (1) CRESST's Java Mapper; (2) a simulated World Wide Web environment; (3) a bookmarking applet; and (4) outcome feedback. In the feedback module, students could access real-time outcome feedback regarding their performance while searching for information and constructing their concept maps. Outcome feedback presented students with information describing whether concept-link-concept propositions in their concepts maps were correct and where additional work was needed. Feedback was based on comparing a student's concept map to that of an expert.

Chang, Sung and Chen's study (2002) addressed that graphic strategies, such as graphic organizers and knowledge maps, had proved useful for text comprehension. In their study, they tested the learning effects of a concept-mapping strategy. Three concept-mapping approaches were designed—map correction, scaffold fading, and map generation—to enhance students' text comprehension and summarization abilities. The experimental results showed that the map-correction method enhanced text comprehension and summarization abilities and that the scaffold-fading method facilitated summarization ability. In addition, Vakilifard, Armand & Baron (2006) examined whether concept maps enhance the second language students' reading comprehension in the first language context. Their result showed that the early presentation of concept maps before reading texts led students in experimental group to understand the texts effectively. Positive results of using concept maps in various subject areas could also be found in Hazzan's (2004), Kwon and Cifuentes' (2007), Ruiz-Primo, Schultz, Li, and Shavelson's (2001) studies.

From the literature and system reviewed, it was found that several problems existed in the studies of mental maps. First, there was a controversial issue regarding the activity using graphic strategies. On one hand, several studies indicated that graphic strategies demanded effort and usually resulted in cognitive overload and negatively affective learning outcomes (e.g., Katayama & Robinson, 2000). Novice students tended to be frustrated in the constructive process of mental maps and unwilling to use this strategy (Fisher, Faletti, Patterson, Thornton, Lipson, & Spring, 1990; Reader & Hammond, 1994). On the other hand, some studies revealed that mental maps served as tools for the reader to build his/her own understanding of a text (Cuevas, Fiore, Bowers, & Salas, 2004; Guerrero, 2003; Payne & Reader, 2006; Tea & Lee, 2004). Readers' construction of a graphic presentation might foster deeper processing. Second, the criteria of evaluating students' mental maps were not clearly revealed in the previous studies. In reading a text, each individual might have very different ways in generating his own concepts. This resulted in the difficulty of comparing experts' "concept maps" with those of students. It also made the evaluation of concept maps questionable.

Third, most findings were based on comparisons of evaluation results obtained immediately after the training was finished (Moore & Readence, 1984). This approach might be good for examining the immediate influences of graphic strategies on text comprehension, but it said little about what happened afterward. Further explorations using longer training duration and more extensive reading materials along with participants of different levels of reading

abilities are worth considering in order to examine the relative effectiveness of different concept-mapping methods. Finally, although some studies have reported performance outcome (e. g. Chang, Sung, & Chen, 2001; Folkesson & Swalander, 2007; Glenberg, Brown, Levin, 2007; Schater et al., 1999), they do not present students' thinking process. A model that demonstrates the relative effects of these processes on outcome performance is needed to answer questions such as: What role does the use of feedback play on subsequent searching and finding relevant textual information?

This study reports on our design of a computer system which assists EFL college students in constructing mental maps of referential resolution in reading texts. In order to reduce the cognitive overload on students, our system offers a list of references for students to select, instead of asking them to figure out the key words or concepts everywhere in a text. The system also provides students with three candidate references as a scaffold when they encounter difficulties in finding out the relationship between two references in drawing their mental maps. Paas (1992) stated that a good way to avoid possible overload on students was to provide scaffolding when they were learning. In addition, the evaluation of students' mental maps in our system is based on referential resolution. Referential resolution provides a clear criterion for the comparisons between experts' maps and those of students. For example, in a short text "I have a sister. Her name is Mary. She is a junior high school student." The mental map of referential resolution is shown in Figure 1.

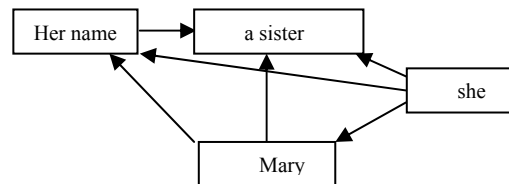


Figure 1. An example of mental map in referential resolution

Furthermore, the examination of three different reading proficiency groups' mental maps of referential resolution also provides a guideline for the teacher to identify individual student's strengths and weaknesses and to plan and design the follow-up remedial courses. Finally, the system records every single reading action that students take in the system and reports these process data back to students and the teacher.

In this study, referential resolution is defined as a reading strategy applied by the reader to interpret the references that have the same meanings as other elements in the text. This operational definition is similar to Paterson, Sanford, Moxey, and Dawydiak's (1998) and Walsh and Johnson-Laird's definition (2004) of co-reference. According to Paterson et al., a major task in reading is to decide whether people or objects appearing in different parts of a text refer to the same entity. While resolving the references, the reader is actually engaged in metacognition which he/she monitors, regulates, and evaluates his/her own reading process (Hartman, 2001). Evaluating one's own reading process involves an assessment of the current progress. This evaluation can assist readers to develop the necessary skills and strategies. Revising one's own reading process refers to modifying previous strategies related to reading goals and other possible learning approaches (Brown, 1987). All these evaluation and revision of one's own reading process are helpful for the meaning construction of text comprehension (Paris & Winograd, 1990).

Coined from the definition of referential resolution, mental representation is identified as the mental map that demonstrates the reader's cognitive structures for establishing the relationships among the references in different parts of a text. In the mental map, references are represented as nodes and connected by lines, which represent the relationships among the references. These lines make the mental map more meaningful by showing how the reader relates references to each other in their cognitive structure. The mental map, thus, is viewed as an external mirror of the internal structure and the process of the reader's thinking (Potelle & Rouet, 2003).

Based on the purpose of assisting EFL college students to practice and construct mental representations of referential resolution in reading, two research questions are addressed: (1) How do the mental maps that students construct in our system help them read and understand the text better? (2) What are the effects of the feedback tool in our system on students' performance of referential resolution?

## System Development

The system built for this study includes three modules, *user interface*, *recording module*, and *feedback module*. The relationships among these three modules are shown in Figure 2. A teacher uses a management interface to save into a database the texts to be read by students. The teacher also prepares the expert's maps that provide the correct reference answers for the texts and must enter the expert maps into the system with the management interface. A student uses a reading interface to work on the exercises by picking references from the texts and construct maps to connect the references. When a student asks for a hint on referential resolution, the system uses the expert's map to provide the student with choices of two distracters and one correct reference. Moreover, every action a student makes in the system is saved into the database by the recording module. The action data provide rich resources from which the teacher can observe many aspects of student behavior during the exercises. Based on the architecture of the system, the interfaces and the modules of referential resolution are shown in Figure 2.

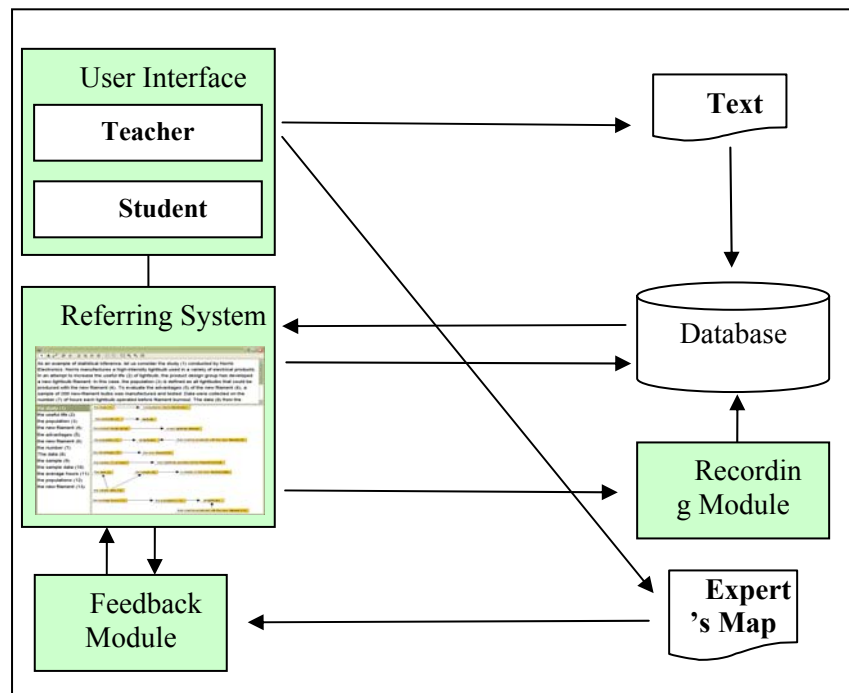


Figure 2. The system architecture

Elements of the student interface are shown in Figure 3. Section A of Figure 3 is a toolbar which includes many graphic tools. *Connection tool* can establish meaningful relations between referential devices. *Feedback tool* compares students' initial map with that of the expert. It then informs students what have been done incorrectly and provides students with three candidate references to correct their previous incorrect selection. Other tools are *cut*, *copy*, *paste*, *erase*, *group*, *ungroup*, *zoom in*, *zoom out*, *undo*, and *redo*. As shown in section B, this area is used to display the text. Students can select a word or a phrase as a text element and then drag into the canvas directly. When a sentence is selected, it will be highlighted.

All referential devices are listed in section C of Figure 3. Students have to understand what these referential devices refer to, then drag and drop them to the canvas. When a referential device in the list is selected, the referential device will also be highlighted in the *text field*. Section D of Figure 3 is a feedback frame. This will inform students what have been done correctly and incorrectly in the referring practice. In addition, the feedback will provide three candidate references when the student has difficulties identifying the link between two words. Students can decide whether they want to activate this frame. Finally, section E is a canvas. On this canvas, students can add links to indicate the relationships between references. They can *add*, *erase*, *drag*, and *drop* elements on the canvas.

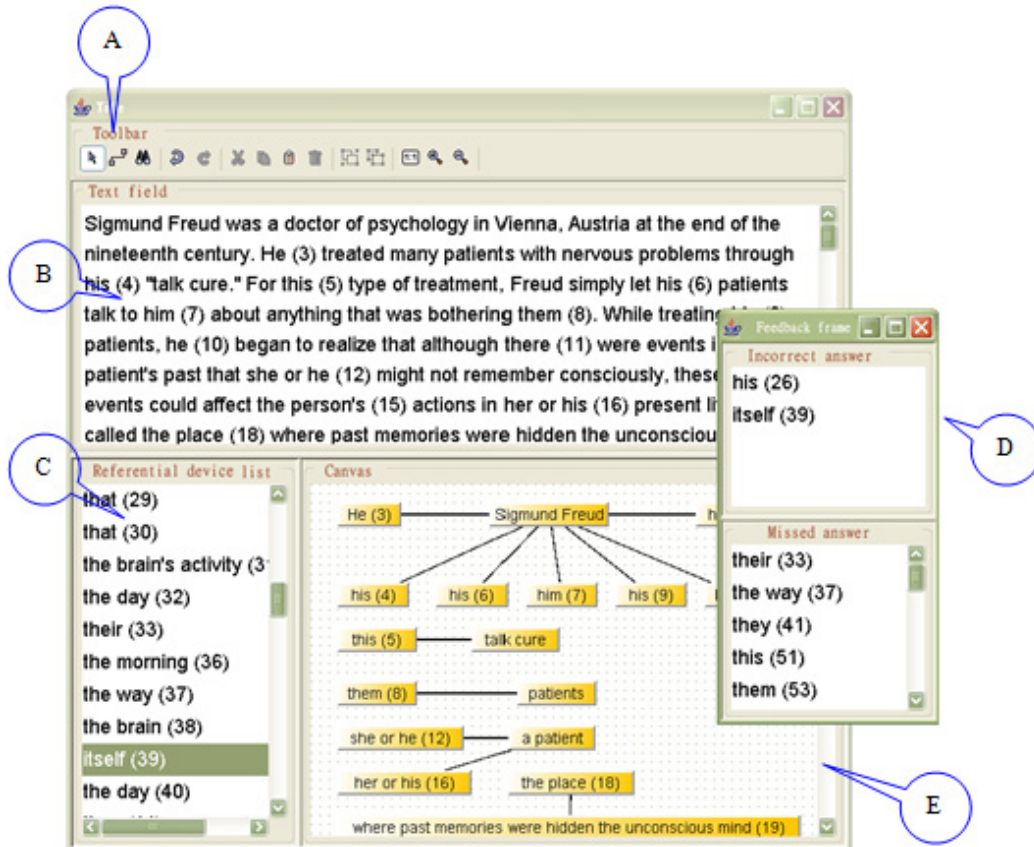


Figure 3. The interface and module of referential resolution

## User Interface

The *user interface* includes a *teacher interface* and a *student interface*. The *teacher interface* helps the teacher manage course data, provide texts which students should read, and observe students' reading process and behavior data. The *student interface* provides a student space to draw a map indicating the relationships between references, take the reading comprehension test, and fill out a questionnaire.

## Recording Module

The system uses a *recording module* to trace students' reading process. From the data, the teacher can identify the difficulties they encounter and check their performance. The records are helpful for the teacher to modify his/her instruction according to students' strengths and weaknesses. The module uses some predicates to record students' behavioral data. Table 1 shows some of the recording predicates.

Table 1. Some of the recording predicates

Predicates	Description
Read a sentence: [T]	Select a sentence T which the student is reading in the text field
Add a cell: [X]	Add an element X on the canvas
Erase a cell: [X]	Erase an element X from the canvas
Add a connection: [X][Y]	Add a meaningful relation between X and Y
Cut a connection: [X][Y]	Cut a relation between X and Y
Get feedback count: [N]	Request the feedback for N times
Get feedback of: [X]	Request the feedback for the reference X

In Table 1, the recording module shows the predicate “read a sentence” when a sentence is selected by a student to read. The predicate “add a connection” is used when the student figures out the relationship between two references and draws a line between these two references. These predicates indicate every single pre-defined action that the student takes in the system. The sequence of these actions then represents the process of referential resolution that the student has gone through in reading a text. An example of *trace result* is shown in Figure 4.

---

1.	Read a sentence: [The young man (22) was hunting with his (23) dogs, and Echo was struck by his (24) perfect figure, beautiful face, and long flowing hair.]
...	
11.	Get feedback count :[26]
12.	Cut a cell :[ the image (51) he (52) saw]
13.	Cut a connect :[the image (51)] connect to [null]
14.	Cut a cell :[his (49) own refecction ]
15.	Add a cell: [refecction]
...	
25.	Get feedback count :[40]

---

Figure 4. Records of a student’s reading process

## Feedback Module

While students encounter difficulties in constructing their initial maps, they can request feedback (see Figure 5). In the example, the student refers “them (8)” to “Sigmund Freud.” He is not sure whether the referential resolution is correct so that he asks for feedback. The system provides the correct reference “patients” and two distracters “that” and “his (9)” and requires the student to pick an answer from the three given choices (see Figure 6).

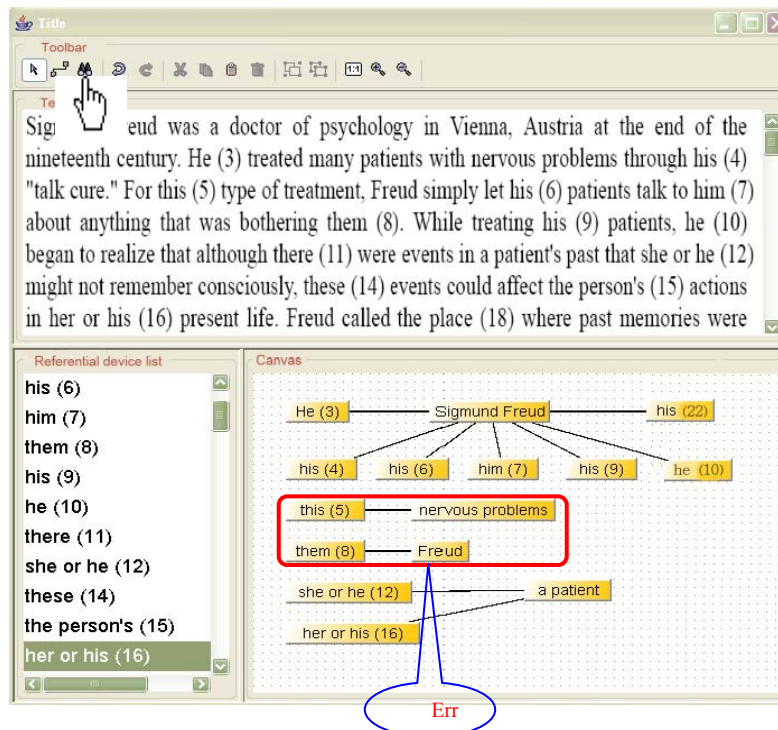


Figure 5. The student asks for a feedback



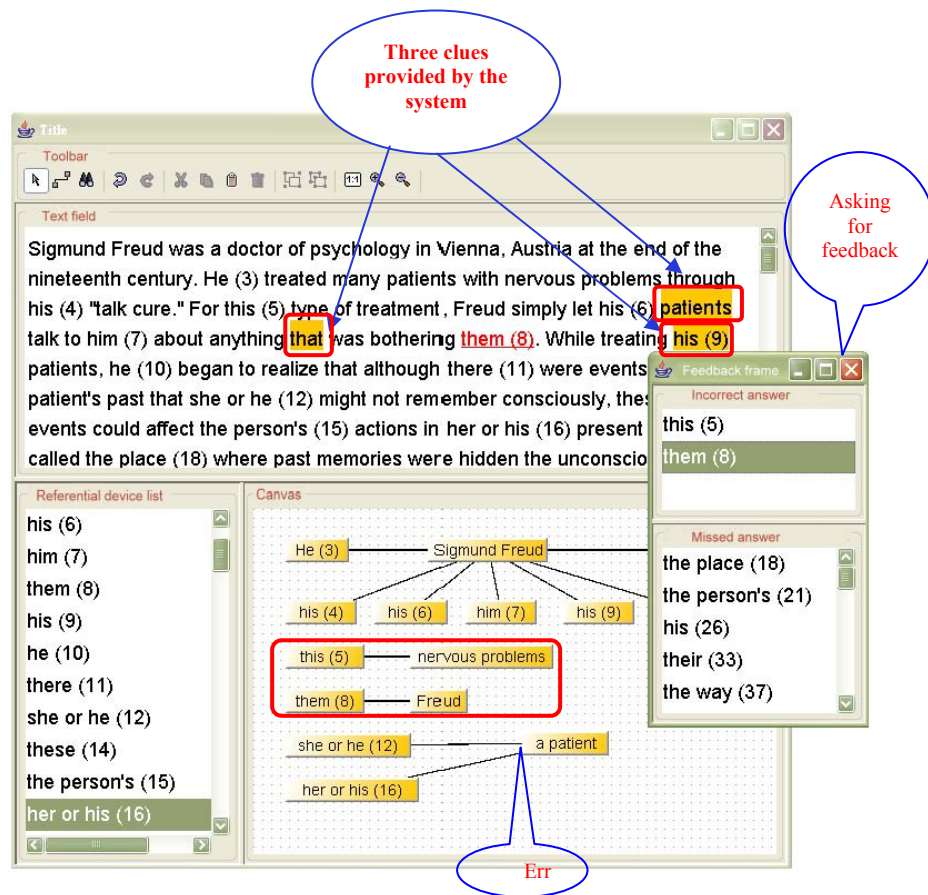
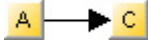



Figure 6. The feedback to a student

A student's answer is compared to the expert's referential resolution in the following way. First, the module transforms the correct map and the student's map into predicates. Second, the module finds all references from students' predicates and then compares them with the correct ones. If students' references do not match the correct ones, they are incorrect. Then the module will provide one correct reference and two distracters as clues back to the student. Table 2 shows the comparison between a student's map and the expert's map in the system. The *feedback module* compares a student's map with that of the expert and finds out that the resolution of the referential device A is incorrect in the student's map. The module will then offer the correct reference B and two distracters D and E as choices for the student.

Table 2. Comparison between a student's map and an expert's map

Student's Map		Expert's Map	
Map	Predicates	Map	Predicates
	Add a cell: [A] Add a cell: [C] Referring: [A][C]		Add a cell: [A] Add a cell: [B] Referring: [A][B]

## Method

### Participants

A total of 90 junior and senior college students were recruited from two reading classes in a technological university in central Taiwan. Their language proficiency levels were defined by their reading scores in a simulated online exam *Testing of English for International Communication* (TOEIC) with a reliability of 0.87. The full score in the online exam was 200.



The frequency distribution of all the participants' score was used to divide the participants into three groups of readers. It was found that the highest frequencies fell in two score intervals, 81-90 (8 students) and 131-140 (8 students). These two intervals were used to identify a less proficient reader group (31-90) and a more proficient reader group (131-200). Participants with scores in the range 91-130 were excluded from the study. The frequency distribution of the participants recruited in the current study is shown in Figure 7.

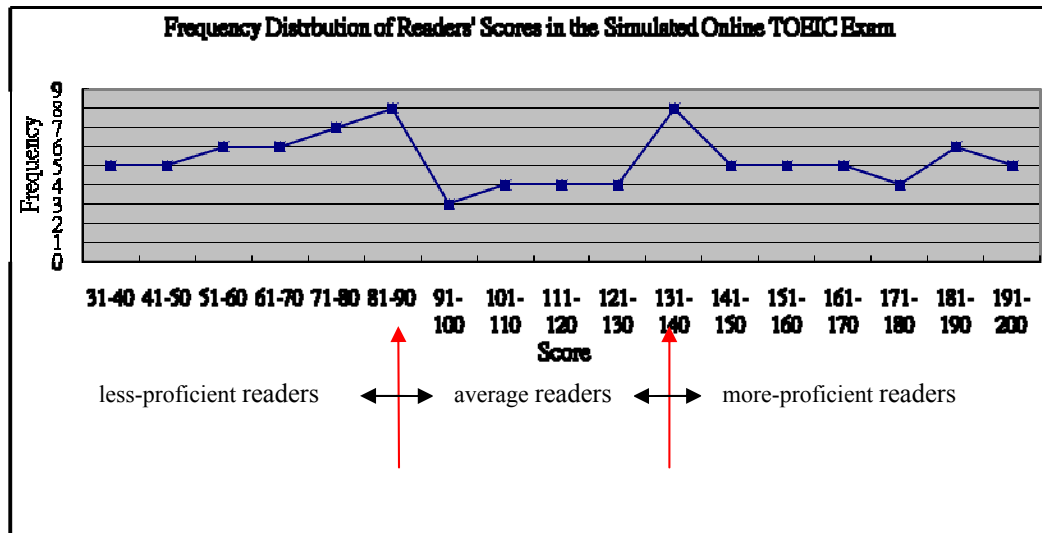


Figure 7. The frequency distribution of the participants' score

Thus, 38 more-proficient readers and 37 less-proficient readers were identified in this study. The mean score of the more-proficient readers was 157.89 with a standard deviation of 18.37. The mean score of the less-proficient readers was 57.84 with a standard deviation of 14.70.

## Material

The online referential resolution practice used three texts to examine the participants' reading comprehension. The three texts were selected from *College Reading Workshop* (Malarcher, 2005) based on the following four criteria: abundance of references for reading practice, similar length, similar readability level, texts written for EFL college students. The three texts were presented in sequence in the textbook and they are text 1---*Freud and the Meaning of Dreams* (number of words: 708; number of referring phrases: 38), text 2---*The Tragedy of Echo and Narcissus* (number of words: 692; number of referring phrases: 62), and text 3---*Commerce through the Internet* (number of words: 651; number of referring phrases: 21). Table 3 shows the readability of these three texts. According to the Flesch reading ease test (Farr, Jenkins, and Paterson, 1951), higher scores indicate material that is easier to read and lower numbers mark harder-to-read passages. The formula for the Flesch reading ease test is shown as follows.

$$206.835 - 1.1015 \left( \frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left( \frac{\text{total syllables}}{\text{total words}} \right)$$

Scores of 90 to 100 are easily understandable by native speakers of English at an average 11-year old students and scores of 60-70 are considered easily understandable by 13- to 15-year old students. Passages with results of 0-30 are best understood by college graduates. In addition, Flesch-Kincaid grade level is calculated with the following formula (Flesch, 1948; Kincaid, Fishburne, Rogers, Chissom, 1975).

$$0.39 \left( \frac{\text{total words}}{\text{total sentences}} \right) + 11.8 \left( \frac{\text{total syllables}}{\text{total words}} \right) - 15.59$$

Among the three texts, text 1 has the most words, 708. Text 2 has the most sentences, 46, and paragraphs, 8. Text 3 is the most difficult with Flesch-Kincaid grade level at 11.5.

Table 3. The readability of text 1 to 3

	Text 1	Text 2	Text 3
Number of words	708	692	651
Number of sentences	35	46	35
Number of paragraphs	7	8	6
Flesch reading ease	53.6	66	44.5
Flesch-Kincaid grade Level	10.6	7.6	11.5

### Procedures of Data Collection

The present study was conducted between April 24<sup>th</sup>, 2006 and June 10<sup>th</sup>, 2006. The 90 college students were asked to complete the online referential resolution practice on the website <http://140.125.32.148/reading> in class. The participants' reading behavior and performance were automatically recorded by trace results in the system.

For each online text, the participants were provided instruction on the definition and types of references followed by examples. They were also asked to practice the referential resolution in the trial section by drawing the relationships between two phrases. After the trials, they were requested to complete a formal referential resolution practice by selecting referential phrases and drawing the relationships between these phrases on the canvas. After drawing the mental maps, they were required to take a reading comprehension test. Each comprehension test contained 8 multiple-choice items. At the end, the participants were requested to fill out an open-ended questionnaire. The online system graded the participants' referential resolution practice by giving one point to each correct connection between two referential words. This was similar to the scoring in the reading comprehension test. Each test item was assigned one point and the full score was 8. The reading comprehension test was specific to each text and an example was shown in the appendix.

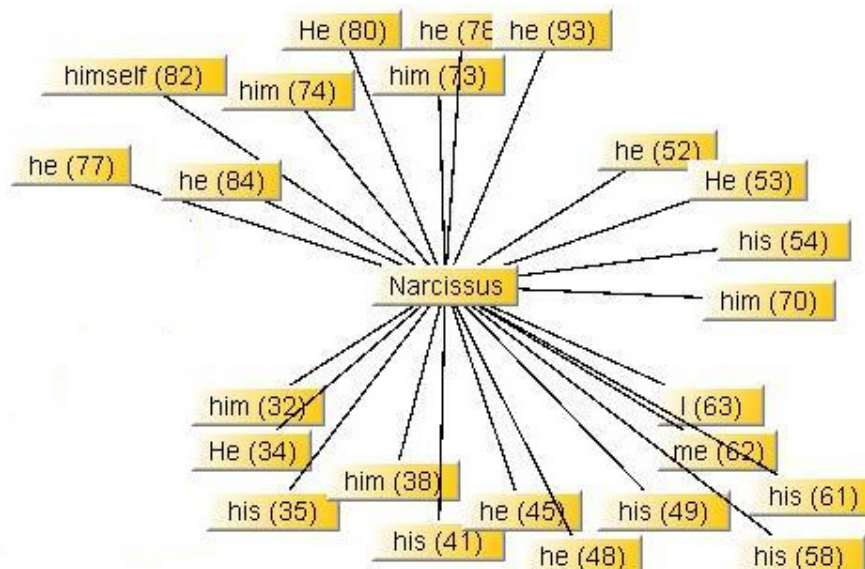


Figure 8. An example of the more-proficient participants' mental map in personal references

### Procedures of Data Analysis

Based on the research questions, the data were categorized into three types: reading process, reading product, and perception towards the referential resolution practice and the online system. Reading process refers to the mental maps of the participants' referential resolution and the trace results of their reading behavior. Reading product includes the participants' scores of referential resolution, scores of reading comprehension test, and frequency of feedback. Finally, students' perception towards the resolution task and the online system will be described.

## Result

### Reading Process

In this study, data from the reading process include the mental maps of the more- and less-proficient participants' referential resolution and the trace results of their reading behavior. An example of the more-proficient reader's mental maps in the text *The Tragedy of Echo and Narcissus* is shown in Figure 8.

As shown in Figure 8, the more-proficient participant was able to integrate personal references in different parts of the text to form a coherent network. The more-proficient participant could refer "his", "him," "himself", and "he" to "Narcissus." He used references from No. 32 to 84. This simple task for the more-proficient participant became a big challenge for a less-proficient participant. The less-proficient participant could not relate references toward the same subject as well as the more-proficient participant did. Figure 9 illustrates an example of the less-proficient participants' mental maps in personal reference. He could only used references from No. 53 to 63. This implies that he could not understand the full text.

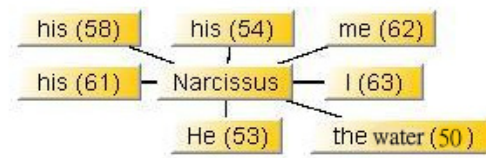


Figure 9. An example of the less-proficient participants' mental map

	1. Add a cell: [Echo]
	...
	26. Add a connection: [she (37) connect to [Echo]
	27. Add a connection: [I (39) connect to [Echo]
	...
	74. Get feedback count :[2]
	75. Cut a connect :[his (23)] connect to [null]
	76. Cut a connect :[his (24)] connect to [null]
	77. Get feedback count :[3]
	78. Cut a connect :[he (26)] connect to [null]
	79. Get feedback count :[4]
	...
	86. Get feedback count :[6]
	87. Get feedback count :[7]
	88. Get feedback count :[8]
rereading	89. Read a sentence: [The young man (22) was hunting with his (23) dogs, and Echo was struck by his (24) perfect figure, beautiful face, and long flowing hair.]
	90. Read a sentence: [The young man (22) was hunting with his (23) dogs, and Echo was struck by his (24) perfect figure, beautiful face, and long flowing hair.]
	...
Self-correction	168. Get feedback count :[26]
	169. Cut a cell :[ the image (51) he (52) saw]
	170. Cut a connect :[the image (51)] connect to [null]
	171. Cut a cell :[his (49) own refection ]
	172. Add a cell: [refection]
	...
Frequency of asking for feedback	259. Get feedback count :[40]
	260. Cut a connect :[the water] connect to [the image (51)]
	261. Cut a cell :[the image (51)]

Figure 10. Trace result of a more-proficient participant's reading process

As shown in Figure 9, the less-proficient participant's referential resolution was more inconsistent than that of the more-proficient participant. Even though he tried hard to resolve the references, his/her answers were incorrect. For example, the less-proficient participant referred "the water" to an incorrect personal reference "Narcissus" in Figure 9.

In addition to the mental maps, the participants' reading process was recorded in the trace results of the online system. Figure 10 is an example of the more-proficient participant's trace results with the text *The Tragedy of Echo and Narcissus*.

The *trace result* showed that the more-proficient participant paid more attention to their reading. He often read and reread the sentence to make sure he really understood or not (e.g., lines 89 and 90). He also used the feedback tool 40 times to help him comprehend the text. After receiving the feedback, he was engaged in comprehension monitoring for self-correction, such as adding or cutting a text element (e.g., lines 169 to line 172). In contrast, the less-proficient participant usually skipped the sentence that he/she did not understand instead of asking for feedback. An example of the less-proficient participants' trace results was shown in Figure 11.

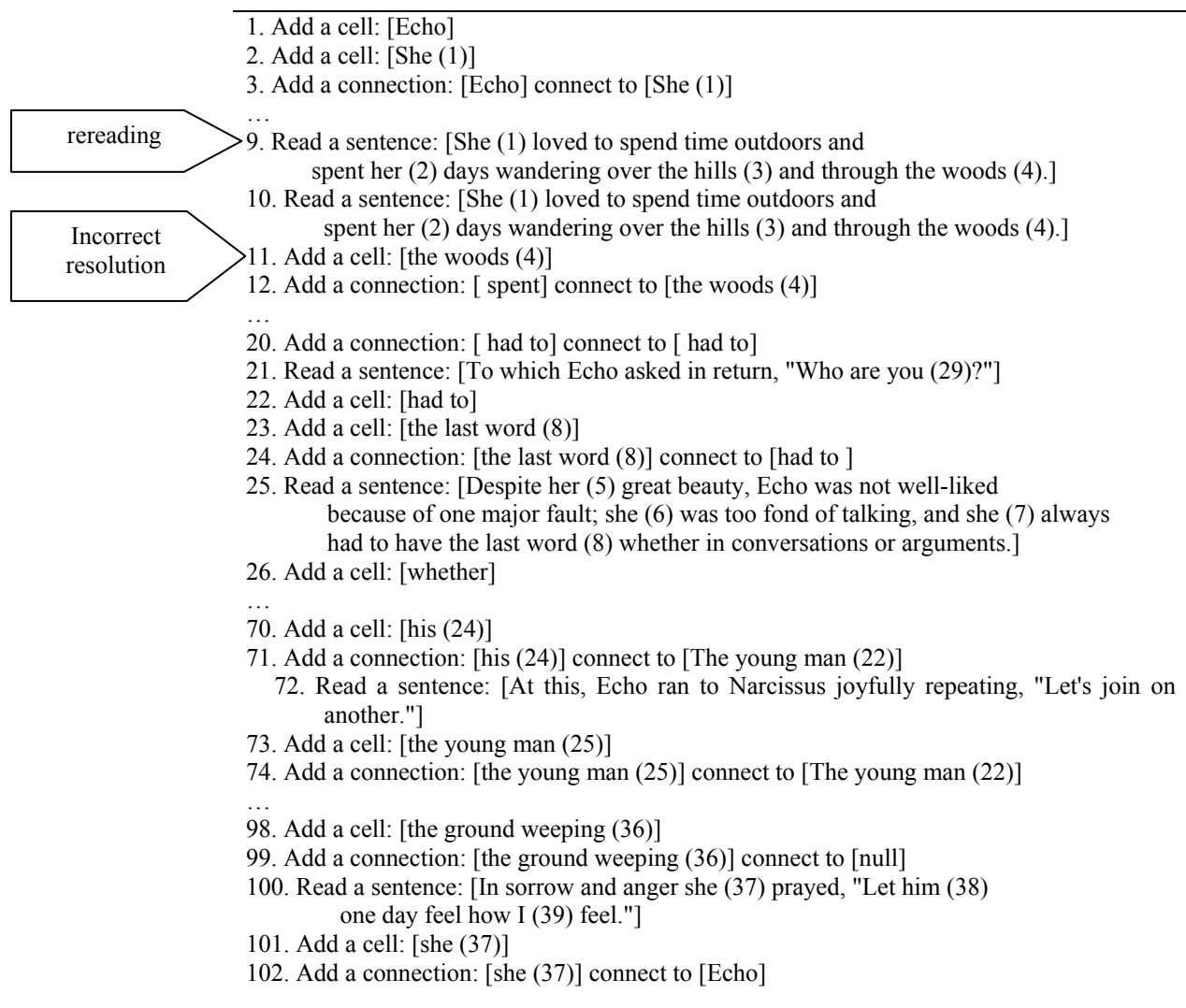


Figure 11. Trace result of a less-proficient participant

Comparing with the more-proficient participant's trace result shown in Figure 10, the less-proficient participant seldom asked for feedback. Although he did self-correction, he did not appropriately revise the incorrect referential resolution. After he read and reread the sentence (e.g., lines 9 and 10), he made a connection between "the woods"

and “spent” (e.g., lines 11 and 12). The difference between the two participants in their comprehension monitoring was that the more-proficient student went through the process of evaluation and revision successfully while the less-proficient student got stuck in these processes.

## Reading Product

The 90 participants’ correct percentage in referential resolution practice and reading comprehension test were shown in Figure 12. The full score of referential resolution practice in text 1 (*Freud and the Meaning of Dreams*) was 38; that of text 2 (*The Tragedy of Echo and Narcissus*) was 62; and that of text 3 (*Commerce Through the Internet*) was 21. As shown in Figure 12, students made progress in the practice of referential resolution as the percentages of correct resolution (the number of correct referential resolution/the total number of references) were 55 % (20.71/38), 70 % (43.42/62), and 69 % (14.47/21) in the three reading tasks. This was similar to their progress in reading comprehension tests as the percentage increased from 61 % (4.85/8) to 82 % (6.55/8) in three texts.

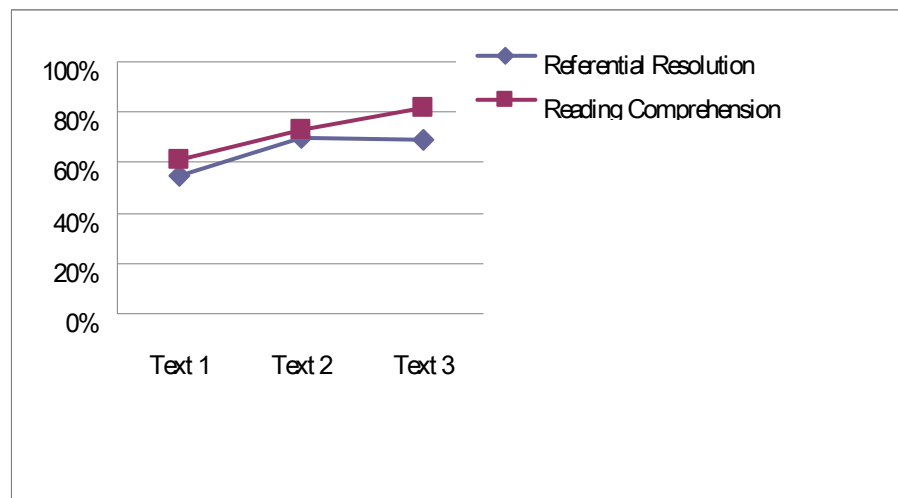


Figure 12. Students’ performance in referential resolution practice and reading comprehension test

As students made progress in referential resolution practice and reading comprehension test, they requested fewer feedbacks as the frequency of feedback dropped from 17.90 in text1, 13.61 in text2, to 7.67 in the final text. This suggested that students became more familiar with the task and more independent in resolving references.

Table 4 presents the result of the Pearson product-moment correlation coefficient between the referential resolution practice and the reading comprehension test.

Table 4. The correlation between referential resolution score and reading comprehension score

Number of participants	Pearson’s correlation		
	Text 1	Text 2	Text 3
90	.61	.67	.75

As shown in Table 4, the referential resolution score and the reading comprehension score had a positive correlation. When the participants got higher scores in referential resolution practice, their scores on reading comprehension test were also higher.

Thirty-eight more-proficient and thirty-seven less-proficient participants’ reading scores in referential resolution practice, frequency of feedback, and reading comprehension test are further shown in Table 5 with mean scores of the more-proficient readers on the left of the slash.

*Table 5.* The mean scores and standard deviation of the more- and less-proficient participants in referring, feedback, and reading comprehension test

	Text 1		Text 2		Text 3	
	mean	SD	mean	SD	mean	SD
Referring	22.30/16	8.39/ 8.45	46.6/32	10.90/ 17.94	16.14/4	3.82/ 3.46
Feedback	16.22/17.63	8.90/ 25.24	14.43/ 10	12.96/ 9.19	7.68/ 2.33	8.44/ .58
Reading	5.86/	1.01/	6.57/	1.21/	7.14/	1.55/
Comprehension	2.63	1.60	4.20	2.17	2.33	.58

According to Table 5, the more-proficient readers outperformed the less-proficient ones in referential resolution practice. Generally, the more-proficient students asked for more feedbacks than the less-proficient students. Both groups decreased their frequency in asking for feedback in three texts. Even though they decreased asking for feedback, they made progress in reading comprehension. That is, students became more independent and less relied on the feedback tool in three sequential reading tasks.

A *t*-Test was further conducted to examine whether there were significant differences between the more-proficient and the less-proficient readers in referential resolution, reading comprehension test, and frequency of feedback. The result showed that  $t(73) = 5.30$ ,  $p < .05$  for the referential resolution practice between two groups. This result was similar to that of frequency of feedback and reading comprehension test. That is, the more- and less-proficient readers' mean scores in the referential resolution practice, frequency of feedback, and reading comprehension test were significantly different.

In addition, the performance of the two groups of participants in referential resolution practice and feedback was evaluated. Table 6 shows the correlation between feedback frequency and the number of errors and between feedback frequency and the number of missed references.

*Table 6.* The correlation between the number of errors and feedback frequency and the correlation between missed referents and feedback frequency

Readers	Text 1		Text 2		Text 3	
	Feedback/ Error	Feedback/ Missed	Feedback/ Error	Feedback/ Missed	Feedback/ Error	Feedback/ Missed
More-proficient readers	-.43	-.38	-.87	-.76	-.50	-.86
Less-proficient readers	.19	-.81	.14	-.62	.36	-.82

According to Table 6, the correlation between feedback frequency and the number of errors was negative for the more-proficient readers. That is, when they asked for more feedbacks, they made fewer errors. This was also true for the relationship between feedback frequency and the number of missed references. In contrast, the correlation between feedback frequency and the number of errors was positive for the less-proficient readers. The feedback tool was not very helpful for the less-proficient students. They did not actively ask for feedback and did not select the correct answers when the feedback provided clues.

## Discussion

The two research questions asked in the introduction section could be answered by the results of this study. First, in the mental representation of referential resolution practice, the more-proficient students were able to construct and integrate references in different parts of the text to form a coherent network while the less-proficient students often referred the references to incorrect or unrelated subjects. The more-proficient students paid more attention to the text elements they read. They often read and reread a sentence to make sure they really understood the sentence. They also used the feedback tool for help. After receiving the feedback, they were engaged in comprehension monitoring for self-correction, such as adding or cutting a text element. In contrast, the less-proficient students usually skipped

the sentence that they did not understand instead of asking for feedback. By examining students' mental representation of textual information and the records of their reading process, the reading teacher can further design appropriate lesson plans and classroom activities to assist the struggling students to overcome their difficulties and compensate their weaknesses. In other words, the mental maps and the *trace results* serve as guides for the reading teacher to design and plan follow-up remedial courses.

Second, students made progress in the online practice of referential resolution as the correctness of resolution increased from 55% to about 69% in three reading tasks. This was similar to their performance in reading comprehension tests as the correctness rate increased from 61% to 82% in the three texts. Further analysis indicated that the correlation between referential resolution and reading comprehension ranged from 61% to 75% in the texts. From these results, it was suggested that referential resolution was one of the essential skills to the enhancement of students' reading comprehension. This online reading system offered students abundant opportunities to practice referential resolution when they tried to understand the text. That is, the online practice of referential resolution could enhance students' reading comprehension. Without referential resolution, students might misunderstand a text to a greater extent.

Third, the *feedback tool* served as a scaffold for students when they encountered difficulties in establishing the relationships between words. It was only provided when the students requested them actively. Since text 1, text 2, and text 3 were presented to students sequentially, a comparison on the feedback frequencies for the three texts could show whether the students relied more, or less on the scaffolding tools as they worked on the texts. The results of this study showed that students gradually asked fewer feedbacks from text 1 to 3 as the mean of feedback frequency decreased from 17.90 to 7.67. In other words, students became more independent and relied less on the feedback in the three reading tasks. This implied that students could use "referential resolution" strategy in the follow-up reading activities.

In an open-ended questionnaire, fifty-one students, 22 more-proficient and 29 less-proficient students, expressed that the referential resolution task was more difficult in text 3 than it was in text 1 and 2. This confirmed the measurement of Flesch-Kincaid grade level of text 3 at the most difficult one, 11.5, among the three texts (see Table 3). In addition to the text difficulty, fourteen students revealed that they still had difficulties in figuring out the relationship between references after using the online system. This deserved more investigation in future study.

The result also indicated that the correlation between the number of errors and feedback frequency was negative for the more-proficient readers. That is, when they asked for more feedback, they made fewer errors. This was also true for the relationship between the number of missed references and feedback frequency. In contrast, the correlation between the number of errors and feedback frequency was positive for the less-proficient readers. They did not actively ask for feedback and seldom selected the correct answer when the feedback provided candidate references. The less-proficient students should be encouraged to ask for feedback when confronting difficulties in referential resolution.

Some limitations were found in this study. First, *trace result* could not reveal all details of the reading process that a reader went through in a text as he might read a sentence without clicking the mouse to select a sentence. This possibly made the process data incomplete in this study. Second, referential resolution could not indicate students' full comprehension of a text. In this study, students' reading comprehension was correlated to their performance on the referential resolution task. Other cohesive ties, such as lexical cohesion, conjunction, substitution, and ellipsis should be used to get a better understanding of students' text comprehension. Finally, the teacher's perception toward students' development of referential resolution should be further explored. An interview could be conducted in the future to investigate the teacher's perception of the impact of the system on students.

Some new help functions could be added to the system. First, after students have completed the referential resolution practice, the system should not only provide the results of grading but also the expert's mental map. This will help students compare their initial mental map with that of the expert. It will also assist them to reflect on their own reading process. The precision rate of the students' mental map should also be provided, so that the students' mental maps in referential resolution can be further analyzed by the students themselves. Through the analysis, students could learn from building connections of related words and make progress in reading comprehension.

Second, the sequential number of references should be removed from the texts and shown in students' mental maps of referential resolution only. This will help students concentrate on reading instead of tracking the numbers of references. Third, the three candidate references that the system provided should be categorized into personal, demonstrative, and locative references. This will help students identify the related references in figuring out their relationships. Finally, the reader-adapted check-points could be taken into consideration for the next implementation of the system. Students could be given more freedom in terms of the type and the number of texts they intend to read based on their specific needs. A reader's self-selected materials could also be adopted in the system that allows readers to set up their reading plan for a whole semester. Another consideration is to use the system for bonus credit purposes to motivate low achievers to use the system, rather than requiring the whole class to complete the same tasks. Individual learning styles and strategies should also be taken into consideration.

## References

- Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading*, 7, 15-32.
- Brown, A. L. (1987). *Metacognition, Executive control, Self-regulation, and Other more mysterious mechanisms*, Hillsdale, NJ: Lawrence Erlbaum.
- Chang, K. E., Sung, Y.T., & Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning*, 17, 21-33.
- Chang, K.E., Sung, Y.T., & Chen, I.D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education*, 71(1), 5-23.
- Chen, S. H. (2001). Cohesion in children's L2 narratives and teaching of discourse skills. *Selected papers from the Tenth International Symposium on English Teaching*, Taipei: Crane Publishing, 310-323.
- Chen, S. Z., & Dai, H. T. (2003). An analysis toward Taiwanese students' cohesive ties in reading comprehension. *Proceedings of The 20<sup>th</sup> International English Teaching and Learning*, Taiwan, 235-258.
- Chu, H. C. J., Swaffar, J., & Charney, D. H. (2002). Cultural representation of rhetorical conventions: the effects on reading recall. *TESOL Quarterly*, 36(4), 511-541.
- Cuevas, H. M., Fiore, S. M., Bowers, C. A., & Salas, E. (2004). Fostering constructive cognitive and metacognitive activity in computer-based complex task training environments. *Computers in Human Behavior*, 20, 225-241.
- Demel, M. C. (1990). The relationship between overall reading comprehension and comprehension of coreferential ties for second language readers of English, *TESOL Quarterly*, 24(2), 267-292.
- Farr, J. N., Jenkins, J. J., & Paterson, D. G. (1951). Simplification of Flesch reading ease formula. *Journal of Applied Psychology*, 35 (5), 333-337.
- Fisher, K. M., Faletti, J., Patterson, H., Thorton, R., Lipson, J., & Spring, C. (1990). Computer-based concept mapping. *Journal of College Science and Teaching*, 19, 347-352.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Folkesson A. M., & Swalander L. (2007). Self-regulated learning through writing on computers: Consequences for reading comprehension. *Computers in Human Behavior*, 23, 2488-2508.
- Fortanet, I. (2004). The use of "we" in university lectures: reference and function. *English for Specific Purposes*, 23(1), 45-66.
- Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, 25 (1), 1-37.
- Glenberg A.M., Brown M, & Levin J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology*, 32, 389-399.
- Greene, J.A., & Azevedo, R. (2007). Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research*, 36(2), 125-148.
- Guerrero, A. M. (2003). *Visualization and Reading Comprehension*, ERIC Document Reproduction Service No. ED 477 161.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*, New York: Longman.



- Hazzan O. (2004). Mental Constructions and construction of web sites: learner and teacher points of view. *British Educational Communications and Technology*, 35 (3), 323-344
- Hartman, H. J. (2001). Developing students metacognitive knowledge and skills. In H. J. Hartman (Ed.), *Metacognition in Learning and Instruction: Theory, Research, and Practice*, Boston: Kluwer Academic, 33-68.
- Huang, S. H. (2005). *Assessing the Relationship between Referential Understanding and Academic Reading Comprehension among EFL College Students*, Masters Thesis, National Yulin University of Science & Technology, Taiwan.
- Katayama, A. D., & Robinson, D. H. (2000). Getting students “partially” involved in note-taking using graphic organizers. *The Journal of Experimental Education*, 68, 119-133.
- Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language*, 49, 335-352.
- Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). *Deviation of new readability formulas for Navy enlisted personnel*, Research Branch Report 8-75, Millington, TN: U. S. Naval Air Station.
- Kwon, S. Y., & Cifuentes, L. (2007). Using computers to individually-generate vs. collaboratively-generate concept maps. *Educational Technology & Society*, 10(4), 269-280.
- Malarcher, C. (2005). *College Reading Workshop* (2<sup>nd</sup> Ed.), Sachse, TX: Compass.
- Moore, D. W., & Readence, J. E. (1984). A quantitative and qualitative review of graphic organizer research. *The Journal of Educational Research*, 78, 11-17.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem solving skill in statistics: a cognitive load approach. *Journal of Educational Psychology*, 84, 429-434.
- Paterson, K. B., Sanford, A. J., Moxey, L. M., & Dawydiak, E. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory and Language*, 39, 290-306.
- Payne, S. J., & Reader, W. R. (2006). Constructing structure maps of multiple on-line texts. *International Journal of Human-Computer Studies*, 64, 461-474.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers’ prior knowledge on the comprehension of hypertext. *International Journal of Human- Computer Studies*, 58 (3), 327-345.
- Pritchard, R. M. O., & Nasr, A. (2004). Improving reading performance among Egyptian engineering students: principles and practice. *English for Specific Purpose*, 23, 425-445.
- Read, W., & Hammond, N. (1994). Computer-based tools to support learning from hypertext: concept mapping tools and beyond. *Computers and Education*, 12, 99-106.
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, 38(2), 260-278.
- Schacter, J., Herl, H. E., Chung, G. K. W. K., Dennis, R. A., & O’Neil, H. F., Jr. (1999). Computer-based performance assessments: a solution to the narrow measurement and reporting of problem-solving. *Computer in Human Behaviors*, 15 (3-4), 403-418.
- Sharp, A. (2003). *Reading Comprehension and Text Organization*, Lewiston, New York: The Edwin Mellen Press.
- Tea, A. J. H., & Lee, B. P. H. (2004). Reference and blending in a computer role-playing game. *Journal of Pragmatics*, 36, 1609-1633.
- Vakilifard, A., Armand, F., & Baron, A. (2006). The effects of ‘concept mapping’ On second language learners’ comprehension of information text. *Proceeding of the Second International Conference on Concept Mapping*, San Jose, Costa Rica.
- Walsh, C. R., & Johnson-Laird, P. N. (2004). Co-reference and reasoning. *Journal of Memory & Cognition*, 32(1), 96-106.
- Wang, Y. F., and Ding, W. Y. (1998). Teaching university students reading through text/discourse analysis. *Proceedings of the 15<sup>th</sup> Conference on English teaching and learning in the republic of china*, Taipei: Crane Publishing, 127-148.
- Yu, K. H. (1993). *The Effects of Textual Coherence on EFL Reading Comprehension*, Taipei: Crane Publishing Co.

## Appendix

Test items in an online reading comprehension test (text 2)

- 1) According to the passage, which is not Echo's characteristics?  
(A) Talkative  
(B) Beautiful  
(C) **Indifferent**  
(D) Sensitive
- 2) Which of the following sentences is true?  
(A) Narcissus was distracted by many women.  
(B) Narcissus eventually fell in love with Echo.  
(C) Narcissus loved Hera before.  
(D) **Narcissus eventually fell in love with himself.**
- 3) Why was Echo attracted by Narcissus?  
(A) **He has charming appearance.**  
(B) He has a warm heart.  
(C) He has mighty power.  
(D) He has special odor.
- 4) How did Narcissus recognize his figure?  
(A) **through water**  
(B) through a glass  
(C) through many women's words  
(D) through Echo's description
- 5) The word **consumed** in the passage has the closest meaning to  
(A) earned  
(B) learnt  
(C) provided  
(D) **spent**
- 6) What will be the best title for this article?  
(A) The honeymoon of Echo and Narcissus  
(B) **The tragedy of Echo and Narcissus**  
(C) The competition of Hera and Echo  
(D) The origin of Narcissus
- 7) The word withered away in the passage has the closest meaning to  
(A) **fade**  
(B) bloomed  
(C) made  
(D) increased
- 8) According to the passage, the story is  
(A) heroic  
(B) martial  
(C) **grievous**  
(D) joyous

# Valuation of Online Continuing Medical Education and Telemedicine in Taiwan

**Fuhmei Wang**

Department of Economics, National Cheng Kung University, Taiwan // fmwang@mail.ncku.edu.tw

## ABSTRACT

Physicians have acknowledged information technology (IT) efficiency and now utilize it in their professional practice and patient management. The benefits of IT within the health care environment has received academic attention, however existing literature currently pertains to limited areas, including the financial effects of telemedicine networks, and both the acceptability of this branch of medical care and the perceptions of patients and clinicians. Studies on welfare valuation of telecommunication health services in an economy are deficient. Based on a welfare concept valuation method, this research provides systematic as well as empirical analysis of telemedicine and finds that online CME plays a key role in enhancing the health care environment. The quantity of health services, the quality of the health care environment, as well as a consumer surplus all increased with online CME lectures. Such lectures are underprovided in telemedicine systems. Appropriate government intervention or programs, through adding physicians for local on-site visits with the electronic health care program and increasing the availability and quality of online CME, could possibly remedy the situation and establish a stable, well-structured, and effective medical care system for rural areas.

## Keywords

Continuing medical education (CME), Telemedicine, Effectiveness, Welfare valuation

## Introduction

Telemedicine is the use of electronic information and communication technologies to provide and support health care when distance separates the participants. Because of hyperspecialisation and multidisciplinary, the technical medical environment has profoundly transformed in recent decades. The area of telemedicine is relevant for the treatment of ill-health and for the provision of medical consultation. Existing wisdom from decision makers and psychiatric consultants evaluates clinical applications of telemedicine, including the quality and accessibility of health care (US Congress, 1990). Two assessments contributed to the evaluation. One focuses on the acceptability of medical care and perceptions of patients and clinicians (Puig-Junoy et. al, 1998). The other concentrates on the financial effect of telemedicine in rural health care networks (Bashshur, 1995).

Research conducted by Berman & Fenaughty (2005) on rural Alaskan health services revealed that such new health care technology may increase patient benefits. Similar concerns about accessibility can be applied to Taiwan, a small heavily-populated land area with many mountains and isolated islands. Chen et al. (2001) performed a brief review of the telemedicine system in Taiwan and found that high quality medical care is difficult to access and medical resources are unequally distributed in rural areas. Due to geographic barriers, several archipelagos are deficient in specialist clinicians and medical resources are under-distributed. Primary healthcare is not emphasized and consulting a family physician is not encouraged. Telemedicine services were introduced in 1995 with a particular focus on providing healthcare in rural areas, continuing medical education (CME) for physicians in these areas, and special medical services for the elderly, the handicapped, and terminal ill patients at home. Online CME, one of the important functions of the telemedicine system, leads to a reduced sense of isolation for physicians and enhances the confidence of medical staff and patients.

Based on technically sound for society concern, Brodie et al. (2000) and Autor (2001) found that access to online services can mean better information about jobs, education, and health. This finding can also be applied to service industries, including health care, manufacturing, transportation, government and education. Health care provisions and relevant medical education through telecommunication is perhaps related to increasing social welfare and the development of new health services. Telemedicine systems offer new ways to practice medicine and “are technologically based innovative systems for the remote delivery of personal health services, continuing medical education, and patient health education.” (Bashshur, 1995, p. 22) This research then investigates the educational,

medical, and economic prosperity of government expenditure in infrastructure and health service in the context of telemedicine diffusion.

There are currently no studies that discuss the welfare concepts of telecommunication health services within the managed care environment or the influence of online CME on such health management. To address the relevant effects of access to care for rural residents, while facilitating the continuation and consolidation of positive effects on service delivery and utilization, this research makes an effort to provide an empirical base for evaluating the current operations of health care provisions, and to assess its potential replicability.

Drawing on literature from the health care and the educational technology fields, we measure and value the relevant environment and consequences from multiple perspectives. A better understanding of concerns from all affected factors will help establish the quantitative significance for possible medical interventions or programs. Investigation results serve as paradigms for selecting areas to maximize welfare in health care provisions.

### **Teleconsultation and Distance Education Network**

At least two sites were connected by a broadband network (e.g. asynchronous transfer mode) or a narrowband network (e.g. integrated service digital network). The network allows two-way or multiple-way video face-to-face teaching, learning, and audio interactive communication in real time between participants. The efficiency of CME on the information technology (IT) promised web site has been reported by Brace-Govan & Gabbott (2004). In the telemedicine context, Chen et al. (1998) have demonstrated that video-on-demand provides web-based distance learning courses and multimedia teaching materials as well as allowing physicians to attend synchronous CME or web courses at home or at the work place to obtain CME credits.

### **Aims of this Study**

This study aims to implement innovative methods of blended learning strategies that include the use of new information technologies in continuing medical education. This study further aims to analyze both the impact of these strategies on social welfare improvements as well as the support the telemedicine system offers rural practitioners and teaching hospital consultants interested in the latest research information. This paper is not the first to investigate the effects of blended learning in a medical education framework. Pereira et al. (2007) and McNulty et al. (2000), among others found that the synergic combination of traditional teaching methods with online learning tactics improved academic performance, long-term knowledge and facilitated active learning. Despite the available literature, studies into the impacts of educational technology on social welfare have not considered the important factor of human-computer interactions. Utilizing our knowledge of the welfare concept valuation approach, this study will attempt to explore this distinctive and new area.

This research aligns the feasibility of an integrated approach with the economical side of an online learning system to continue medical education. The methodological goal of this study is to highlight the welfare feature of the telemedicine system and to investigate whether health services increase with online CME lectures for physicians through a telemedicine system, and to test the influences of such increases on welfare variations. Although our findings focus specifically on locations where geographical limitations may impede access to public sector health services for rural residents, our research may be equally relevant for other countries who wish to utilize an electronic health care system to improve access to health care for people living far away from medical centers with and without health insurance.

### **Definitions of Relevant Terms**

In this study, relevant terms are compiled from the literature, technological products in education, production of learning environments, instructions of healthcare provisions, and economics usages. Various terms are defined in order of appearance.

*Continuing medical education* (CME) is required to maintain and further develop competence for continuing professional development to support rural practitioners as well as teaching hospital physicians and consultants who are seeking the latest research information. *Utility* means satisfaction from consuming goods or services. The *demand curve* is generally assumed that the lower the price of goods, the greater the quantity demanded. *Consumer surplus* is presented as the difference between an agent's willingness to pay for goods and actually purchasing value of the goods. It is regarded as an increase in happiness or utilities. *Welfare variation* is usually presented by a change in the consumer surplus. *Effectiveness* analysis guides the choice for the program with the lowest cost to supply fixed units of effects.

## Method

The blended learning method is generally recognized as a supporting system to off-line teaching and learning. Blended learning combines personal interaction from live class sessions with online education for greater learning flexibility (Demetriadis & Pombortsis, 2007). This study investigates the potential merits of using appropriate technology (e.g. computers combined with satellite links/wireless hotspots) to assist in fostering a sense of belonging to one supportive learning community among distance physicians (educators and learners) and improve educational, medical, and social outcomes. It presents supportive networks and technological perspectives about online continuing medical education in the healthcare provision system, while still maintaining its statistical rigor. Furthermore, based on a welfare concept valuation method approach, this study uses the demand theory to measure the value of medical care and with STATA 9.0, produces estimates based on three areas: (1) the quantity of medical services demanded of telemedicine, (2) the welfare status of a given quantity of such services, and (3) possible welfare variation, resulting from an increase in online CME lectures.

## Subjects

The subject of this study is to establish a systematic framework to meet the specific health care demands in offshore archipelagoes in Taiwan. Based on the experience of rural Alaskan health services, Berman & Fenaughty (2005) revealed that such new health care technology may improve treatment for patients. Similar concerns about accessibility can be applied to Taiwan, a small heavily-populated land area with many mountains and isolated islands. High quality medical care is difficult to access and medical resources are unequally distributed in rural areas. Due to geographic barriers, several archipelagos are deficient in specialist clinicians and medical resources are under-distributed. Primary healthcare is not emphasized and consulting a family physician is not yet encouraged. Telemedicine service was introduced in 1995 particularly to facilitate medical care in remote areas and to assist physicians located in these areas in continuing their online medical education. Telemedicine Service further offers special medical services for the elderly, the handicapped and terminally ill patients at home.

Based on the systematic framework, regression analysis is undertaken. The sample covers the period from 1995-2004. Five teleconsultation medical centers, all funded by the Department of Health were surveyed, including National Taiwan University Hospital, Veterans General Hospital Taipei, National Cheng Kung University Hospital, Tri-Service General Hospital, and Buddhist Tsu-Chi General Hospital. This is also regarded as a national sample.

## Systematic Framework

Economic assessment provides a systematic framework to make decisions regarding known costs and the positive effects on health care. Drummond et al. (1997) indicated that economic evaluation should offer an analysis of courses of action in terms of costs and consequences, including direct, indirect and intangible factors. Bashshur (1995) precisely defined the telemedicine system and offered guidelines for the successful assessment of such a regime. Brent (2003) was sure that all resource allocation should provide the highest satisfaction. Henderson (2005) regarded improving the health of a given population as the primary goal of the health policy and believed that the preferred measure of health benefits, being the health outcomes themselves, should not be limited to a dollar value. This research, inspired by Bashshur (1995), intends to assign values to human health and puts the analysis on the willingness-to-pay basis that is so fruitful in other contexts (e. g. environmental studies). The framework examines

the central result of the welfare status responding to changes in the quantities of continuing medical education (CME) lectures through the educational technology system.

No matter how ingenious the investigator, some benefits and costs are impossible to value. The tools of cost-benefit analysis can be used to force planners to reveal limits on how they value intangibles. A program is admissible if the measured costs are less than the measured benefits. Using the tools of demand theory, we investigated the incremental (marginal) value consumers attach to additional consumption of medical care at any level of consumption observed. This is the willingness-to-pay interpretation of a demand curve. As mentioned by Phelps (2003), willingness-to-pay slopes downward for two reasons: (1) diminishing marginal productivity of health care and (2) the decreasing marginal utility of health care itself in producing utility. The key lesson from this work is that individual patients' preferences differ regarding the various aspects of health, and to assume otherwise leads to a loss in utility. Health is composed of many measurable components. The maximizing utility problem is written as:

$$\max_{T,O} U(O, T, Q) \quad (1)$$

$$\text{s.t. } I = O + RT \quad (2)$$

Where  $O$  is the unit price of goods other than medical care, being normalized to unity,  $T$  is the quantity of consumption on telemedicine health services,  $Q$  is the quality of health care environment,  $I$  is personal income, and  $R$  is the monetized cost for accessing the medical services. Consumer objectives are to maximize their utility from consuming telemedicine health care and other goods while being protected by the quality of an electronic health care environment, presented by equation (1), but limited by income and by prices, presented by equation (2). Solving equations (1) and (2), we express the demand function for telemedicine of the representative consumer as:

$$T^* = T^*(R, I, Q) \quad (3)$$

The demand curve is depicted in Figure 1. A consumer surplus is the value of an electronic health service minus the price paid for it, summed over the quantity bought and supplied. The area is equal to the shaded area, the base of the triangle multiplied by the height of the triangle divided by 2.

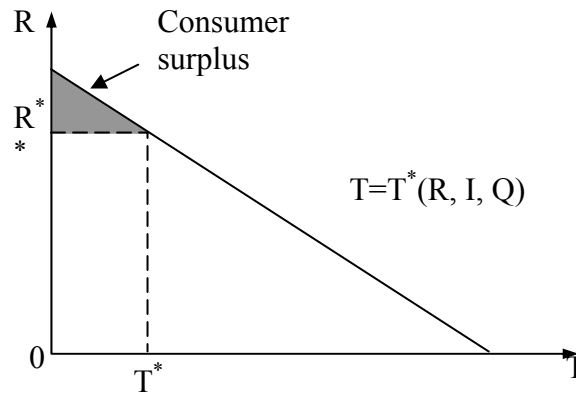


Figure 1. The demand curve and consumer surplus of telemedicine health services

## Empirical Tests

After constructing the research model, we conducted an empirical test on the appropriate allocation of health care expenditure to provide health services in accordance with the demands of residents. Our primary interest is whether the health care system of telemedicine leads to consumer surplus and what factors bring welfare improvement in rural Taiwan.

## Specifications and Data Descriptions

From the sustainability evaluation of the demonstrated program, the demand function for health care through the development of telecommunication system is specified as:

$$Quantity = f(Tel. Exp. per service, GDP per capita, NHE per capita, OLCME lectures) \quad (4)$$

As the demand theory demonstrated, the quantity demanded for health services is the number of times a patient plans to access the health care during a given time period at a particular price. In Taiwan, health services are provided by the consolidated health sector in the National Health Insurance system which began in 1995. The quantity of health services (*Quantity*) and telemedicine expenditure per health service (*Tel. Exp. per service*) respectively present the quantity demanded and price in such a regime. Following Bashshur (1995), when optimizing the effectiveness and efficiency of telemedicine, online CME is suggested. From empirical specifications, this research attempts to examine whether online CME plays a key role in minimizing incremental costs per unit of incremental health effect, defined as effectiveness, and providing the medical services that we value mostly highly, defined as efficiency in the telemedicine system.

Table 1. Summary statistics

Variables	Observations	Average	Std. Deviation	Minimum	Maximum
<i>Quantity</i>	10	954.90	437.07	426	1161
<i>Tel. Exp. per service</i>	10	725.32	345.64	296.72	1368.05
<i>GDP per capita</i>	10	13545	692.97	12488	14633
<i>NHE per capita</i>	10	730.55	63.56	636.36	848.00
<i>OLCME lectures</i>	10	1344.5	1221.05	100	3066

Table 2. Correlations among variables

Variables	<i>Quantity</i>	<i>Tel. Exp. per service</i>	<i>GDP per capita</i>	<i>NHE per capita</i>	<i>OLCME lectures</i>
<i>Quantity</i>	1				
<i>Tel. Exp. per service</i>	-0.8	1			
<i>GDP per capita</i>	0.62	-0.38	1		
<i>NHE per capita</i>	0.69	-0.48	0.73	1	
<i>OLCME lectures</i>	0.55	-0.33	0.16	0.74	1

Quality indices of the telemedicine health care environment, represented by effectiveness and efficiency, involve more than one indicator. The quality of the health care environment in rural Taiwan is composed of national health expenditure per capita (*NHE per capita*), provided by the consolidated health sector, and online CME lectures (*OLCME lectures*). The data of the above variables are taken from *Bureau of Nursing and Health Service Department and Bureau of National Health Insurance*, Taiwan (2005). The explanatory variable *GDP per capita*, published by the *Council for Economic Planning and Development*, Taiwan (2005), presents the income status of the consumers. This indicator is verified by Pritchett & Summer (1996) and Olsen & Dahl (2007), indicating that GDP per capita is strongly associated with better health in a richer economy. The sample covers the period of 1995-2004 and 5 medical centers founded by the Department of Health in the telemedicine system. Tables 1 and 2 respectively display the summary statistics and the correlation matrix between variables of the implementation for the entire period.

## Regression Results

Regressions of medical services on selected variables are presented in Table 3. Empirical evidence from Taiwan experiences conveys a wealth of information. The coefficient on telemedicine expenditure per health service (*Tel. Exp. per service*) indicates that raising the cost of per service in telemedicine significantly decreases the quantity demanded of such health care. An increase in the cost of 0.10 reduces the quantity of services to around 0.08. A 0.1 increase in GDP per capita significantly increases the health services by 0.06. Increasing NHE per capita is associated with lower medical care quantities, whereas increasing online CME lectures for physicians (*OLCME lectures*) leads to a rise in the quantity of health services.

Table 3. Regressions of quantity in telemedicine health services on selected variables

Number of observations=10 F(4, 5)=14.43 Prob>F=0.0059 R-squared=0.92 Adjusted R-squared=0.86 Root MSE =165.56						
Quantity	Coef.	Std.Err.	t	P> t	95% Conf. Interval	
<i>Tel. Exp. per service</i>	-0.76	.18	-4.15	0.009***	-1.23	-0.29
<i>GDP per capita</i>	0.61	0.22	3.07	0.028**	0.11	1.26
<i>NHE per capita</i>	-8.10	3.61	-2.24	0.075*	-17.38	1.19
<i>OLCME lectures</i>	0.38	0.13	2.94	0.032**	0.05	0.71

\*: significant at 10%; \*\*: significant at 5%; \*\*\*: significant at 1%

In Taiwan, under government support, most of the medical centers participate in the integrated delivery system and dispatch additional specialist clinicians. Air-transport transfers are adopted for the remote delivery of health services, increasing with the level of NHE per capita, and in turn decreasing with outpatients through telemedicine systems. However, to enhance the confidence of medical staff and patients and to reduce a sense of isolation for physicians, the archipelagos in Taiwan are covered by the telemedicine program for medical consultation and OLCME. The quality of the health care environment through telemedicine increases with online lectures of CME for physicians.

The positive relationship between health services, the health of a nation and its economic prosperity is well recognized (Pritchett & Summers, 1996; Olsen & Dahl, 2007). Patients in a telemedicine system are positively affected by the level of GDP per capita in the economy. Our primary interest is whether the adoption of a telemedicine system improves the welfare status of the economy. Ruling out trivial cases and using the demand curve to measure value of such care, influences of environment and income status in the systematic framework are then averagely specified in the constant term. The demand curve to be estimated is:

$$Quantity = 2855.9 - 0.76 Tel. Exp. per service \quad (5)$$

Substituting the means value of *Tel. Exp. per service*, 954.9 from statistics presented by Table 1, into equation (5), the demand curve for telemedicine health care is depicted by Figure 2. In addition, we predict that adopting a telemedicine system leads to a consumer surplus of US\$777517.75, as indicated by the triangle *abc*. The measured benefit or willingness-to-pay is greater than the measured cost or telemedicine expenditure per service. According to the guideline of cost-benefit analysis, the program is admissible. The consumer surplus increases with the quantity of health care services provided by the telemedicine system.

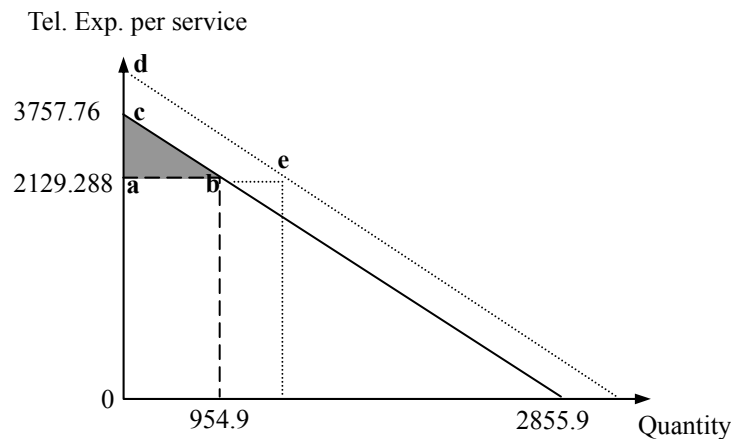


Figure 2. A welfare concept valuation of telemedicine systems



## Discussions

Increasing online CME lectures, assuming no change in other factors, shifts the demand curve outwards to account for the improvement in the quality of medical care. These increases led to a reduction in feelings of isolation for physicians and enhanced the confidence of medical staff and patients. A sizeable amount of transfers can then be reduced and through the telemedicine system, health services will increase and improve. As a result, the consumer surplus in the economy rises, which increases by rectangle *bcd*e in Figure 2. The implication is clear: the system underprovides online CME lectures. Appropriate intervention or program could possibly remedy this situation. Of course, all the difficulties in measuring the quantity and value of the health care still remain.

In the same spirit, there have been many studies on the value of a life. Viscusi & Aldy (2003) estimate the value of a life is between \$4 million and \$10 million. When considering this value and considering the difficulties in putting a price on saving lives, the quality of health care currently available in rural areas and online CME lectures offered to rural health-care staff are deficient and inadequate. Furthermore, the telemedicine program resulted in a reduction in the time health-care personnel devoted to traveling to their referral centre. The monetary value of the saved time is important. From the restricted budgetary perspective, the results demonstrated that the benefits from welfare improvement by increased expenditure introduced by the telemedicine system are bigger than the costs from an increase in the direct national health expenditure within the traditional health-care network, as the fourth row of Table 3 indicates.

The interconnectedness is noticeably being adopted and improved in recent international efforts in healthcare reform. In the future, telemedicine systems are expected to connect hospitals, insurance organizations, public health administrations, and finally, to extend to every household. The satisfaction of health-care staff and households with online educational programs and medical consultations seems to be assured.

Despite the strengths, like all tools it must be used appropriately. Pallof & Pratt (2001, p. 26) emphasized that effective collaboration can enhance the learning experience and can therefore be considered as one of the determining factors in measuring the success and quality of any online course. Kidney & Puckett (2003) has highlighted the presence of problems similar to those detected in the application of e-learning strategies in distance learning courses. The principle aim of the application of telecommunication techniques is to decrease clinicians' effort and time in continuing medical education. From our analytical results, increases in CME lectures raise the health care environment. If so, that would be another, possibly much more important, source of welfare improvement derived from a policy to promote telemedicine.

Hazemi & Hailes (2002) indicated that educational institutions are now compelled to integrate technological and electronic pedagogy into their traditional education product line. Information and communication technologies in other areas of educational study received our attention. The impact of modern technology on the advancement of telemedicine depends on the works of different bandwidths for tele-consultation and distance education programs. With the appropriate technology (e.g. computers combined with satellite links or wireless hotspots), continuing professional development is required to maintain and further develop competence. Rural practitioners and teaching hospital consultants use this technology to carry out the latest research information and transmit the knowledge to where it is needed. However, the inability of web-based systems to fully deal with learning diversity and the high cost of developing new systems might weaken the effectiveness of blended learning in an operational capacity (Nasseh, 2000). This paper is limited to the analysis imposed by assuming that the effectiveness of asynchronous distance education is exogenous. This limitation is not intractable, but requires the use of appropriate surveys to ensure that the results we obtained are not influenced by improper assumptions of exogenous effective conduction in analytical and empirical examinations.

## Conclusions

Physicians have acknowledged information technology (IT) efficiency and now utilize it in their professional practice and patient management. The benefits of IT within the health care environment has received academic attention, however existing literature currently pertains to limited areas, including the financial effects of telemedicine networks, and both the acceptability of this branch of medical care and the perceptions of patients and clinicians. Studies on welfare valuation of telecommunication health services with online learning systems in an

economy are deficient. Based on a welfare concept valuation method approach, this research provides positive analysis of telemedicine.

Many factors, including NHE per capita, GDP per capita, online CME lectures, and telemedicine expenditure per health service among others affect the demand for health services through telemedicine systems. The cost per service strongly and negatively affects the quantity demanded. The demand curve in the price and quantity space slopes negatively. Rising NHE per capita always results from a lack of confidence felt by patients or practitioners requesting air-transport transfers. The relationship between electronic health services and NHE per capita is negative. GDP per capita is significantly associated with more health services, including telemedicine and other medical care. Online CME plays a key role in enhancing the electronic health care environment. Quantity of health services, quality of health care environment, and consumer surplus increase with online CME lectures. This may draw attention to any issues that might need to be considered when introducing online CME, though the sample is limited to ten years. Furthermore, this regime can be designed not only for improving the quality of consultation and referral in the rural areas but also strengthening the confidence of patients and practitioners, both CME and referral. Then hospitals can concentrate on its core missions, providing inpatient services.

The use of multimedia in educational contexts is based on the assumption that interactive learning provides greater educational benefits. This study applies traditional teaching and computer-assisted learning to provide evidence of how teaching hospital consultants approach healthcare policy making, development, and evaluation as they relate to the quality assurance of continuing medical education and learning management systems. Based on educational issues related to quality assurance, a welfare evaluation approach reveals that they tend to have better enhancement in the electronic health care environment. The investigation has guided physicians, medical educators, patients, and practitioners to address key areas in order to reliably assure the quality of consultation and referral in the rural areas supported by learning management systems.

We linked the quality of a telemedicine environment with effectiveness and efficiency, which is one of Bashshur's (1995) key concerns. The increase in online CME lectures efficiently allocates medical resources while raising consumer surplus. It further serves to improve the health condition of rural residents by raising incremental health services without increasing the cost per health service. These findings suggest that for improving the quality of the health care environment for rural residents, the system currently underprovides online CME lectures. Appropriate government interventions or programs through adding physicians for local on-site visits with telemedicine program and online CME lectures can remedy this situation. Establishing a stable, well-structured, and effective medical care system for rural areas is anticipated. Potential replicability of relevant evaluation is applicable for other countries to improve access to health care for people living in remote areas.

## References

- Autor, D. (2001). Wiring the labor market. *The Journal of Economic Perspectives*, 15 (1), 25-40.
- Bashshur, R. L. (1995). On the definition and evaluation of telemedicine. *Telemedicine Journal*, 1 (1), 19-30.
- Berman, M., & Fenaughty, A. (2005). Technology and managed care: patient benefits of telemedicine in a rural health care network. *Health Economics*, 14 (6), 559-573.
- Brace-Govan, J. & Gabbott, M. (2004). General practitioners and online continuing professional education: projected understandings. *Educational Technology & Society*, 7 (1), 51-62.
- Brent, R. J. (2003). *Cost-Benefit Analysis and Health Care Evaluations*, Northampton, MA: Edward Elgar.
- Brodie, M., Flournoy, R. E., Altman, D. E., Blendon, R. J., Benson, J. M., & Rosenbaum, M. D. (2000). Health information, the internet, and the digital divide. *Health Affairs*, 19 (6), 255-265.
- Bureau of National Health Insurance (2005). *Health and Maintenance Expenditure Statistical Data Book*, Taiwan: Bureau of Nursing and Health Service Department and Bureau of National Health Insurance.
- Chen H. S., Guo, F. R., Liu, C.T., Lee, Y.J., Chen, J. H., Lin, C.C., Hou, S.M., & Hsieh, B.S. (1998). Integrated medical informatics with small group teaching in medical education. *International Journal of Medical Informatics*, 50 (1-3), 59-68

- Chen, H. S., Guo, F. R., Chen, C.Y., Chen, J.H., & Kuo, T. S. (2001). Review of telemedicine projects in Taiwan. *International Journal of Medical Informatics*, 61 (2-3), 117-129.
- Council for Economic Planning and Development, (2005). *Taiwan Statistical Data Book*, Taiwan: Council for Economic Planning and Development.
- Demetriadis, S. & Pombortsis, A. (2007). E-lectures for flexible learning: a study on their learning efficiency. *Educational Technology & Society*, 10 (2), 147-157.
- Drummond, M. F., O'Brien, B. J., Stoddart, G. L., & Torrance, G. W. (1997). *Methods of Economic Evaluation of Health Care Programmes*, Oxford: Oxford University Press.
- Hazemi, R., & Hailes, S. (2002). *The Digital University: Building a Learning Community*, New York: Springer.
- Henderson, J. W. (2005). *Health Economics and Policy with Economic Applications*, Australia: Thomson.
- Kidney, G. W., & Puckett, E. G. (2003). Rediscovering first principles through online learning. *Quarterly Review of Distance of Education*, 4 (3), 203-212.
- McNulty, J.A., Halama, J., Dauzvardis, M. F., & Espiritu, B. (2000). Evaluation of web-based computer-aided instruction in a basic science course. *Academic Medicine*, 75 (1), 59-65.
- Nasseh, B. (2000). Forces of change: the emergence of a knowledge society and new generations of learners. In Petrides, L. A. (Ed.), *Case Studies on Information Technology in Higher Education: Implications for Policy and Practice*, Hershey, PA: Idea Group, 217-227.
- Olsen, K. M. & Dahl, S. A. (2007). Health differences between European countries. *Social Science and Medicine*, 64 (8), 1665-1678.
- Pallof, R. M. & Pratt, K (2001). *Lessons from the cyberspace classroom: The realities of online teaching*, San Francisco: Jossey-Bass.
- Pereira, J. A., Pleguezuelos, E., Merí, A., Molina-Ros, A., Molina-Toma's, M. C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education*, 41 (2), 189-195.
- Phelps, C. E. (2003). *Health Economics*, New York: Addison Wesley.
- Pritchett, L. O. & Summers, L. H. (1996). Wealthier is healthier. *Journal of Human Resources*, 31 (4), 841-868.
- Puig-Junoy, J., Saez, M., & Martinez, G. E. (1998). Why do patients prefer hospital emergency visits? A nested multinomial logit analysis for patient-initiated contacts. *Health Care Management Science*, 1 (1), 39-52.
- US Congress, (1990). *Office of Technology Assessment, Recombinant Erythropoietin: Payment Options for Medicare*, Washington: US Congress.
- Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27 (1), 55-76.

## Supporting “Learning by Design” Activities Using Group Blogs

**Georgios Fessakis, Konstantinos Tatsis and Angelique Dimitracopoulou**

Learning Technology & Educational Engineering Laboratory, University of the Aegean, Greece //Tel +30 22 4109 9151 // Fax: +30 22 4109 9175 // gfesakis@rhodes.aegean.gr // tatsis@aegean.gr // adimitr@aegean.gr

### ABSTRACT

The paper presents a case study of the educational exploitation of group blogging for the implementation of a “learning by design” activity. More specifically, a group of students used a blog as a communication and information management tool in the University course of ICT-enhanced Geometry learning activities. The analysis of the designed learning activities, the blog content and log files, as well as the points of view of the students (via a questionnaire and a group interview) expressed upon completion of the activity gives significant findings supporting the researchers’ initial hypotheses about the potential of blogs’ educational applications. Blogs combined with a proper pedagogical approach such as ‘learning by design’ enable teachers to offer high quality learning experiences to their students.

### Keywords

Blogs, Learning by design, e-Learning, Interaction analysis, Higher education

### Introduction

The development and availability of social software applications sets new challenges and opportunities for the learning technology community. The term “social software” refers to computer network software applications supporting groups of actors in communication and interaction (Allen, 2004). According to Allen’s historical record the term “social software” appears sporadically with the first groupware applications during 1990-1999. The term social software is systematically used from 2002 in order to include applications such as forums, wikis, blogs, online multi-user games etc., which are not usually covered by the business oriented term of “groupware”.

The educational community’s interest for social software is mainly based on the idea that it is consistent with the modern learning theories. This is especially the case for constructionism (Papert, 1993) and social constructivism theory (Ernest, 1994; Kim, 2001) that emphasize the importance of learners’ active participation to social activities during learning, which is mediated by the design and construction of artefacts. This growing interest is reflected also on the increasing number of research papers concerning the investigation of social software exploitation for educational purposes. In this paper the authors focus on the use of blogs for the improvement of teaching Didactics of Mathematics to pre-service teachers for young children.

A general review and introduction to the educational uses of blogs is presented by Downes (2004). Downes proposes, among other things, the view that students participating in blogging have opportunities to a) reflect on their texts, b) engage in writing for significant time intervals, and c) trigger long dialogue with their readers leading to new writing cycles. According to the authors’ best knowledge the relevant literature includes a rather small number of published studies concerning Mathematics education training using social software and an even smaller number of works using blogging. For example, Marcelo & Bairral (2007) describe the case study of a community of practice in which future Mathematics teachers interact using an electronic communication environment. The members of the particular community were supplied with content (html pages, java applets, URL directories etc.) as well as communication tools including email, forum and chat. One of the main goals of the researchers was the analysis of students’ interactions in the electronic environment. The analysis of those interactions showed that the students formed a system significantly different from an ordinary community of practice, mainly due to the teachers’ presence and the non-voluntary participation of the students. As a consequence, the researchers propose interesting additional analysis dimensions other than the meaning, community, identity and practice which constitute the basic components of communities of practice (Wenger, 1998). The proposed set of dimensions includes the impact of the work assigned to students, the teachers’ role, the participants’ behaviour and intentions, the dialogic views of the implemented communication etc. In a second reading of Marcelo & Bairral’s work it is possible to see the need for a focused study of the interaction between the learning environment’s distinct characteristics and the pedagogical approach (instructional strategy). In other words, it appears that the electronic learning environments patterns of use are not neutral to the genre of the assigned work.

In Makri & Kynigos (2007) a case study is described concerning the integration of group blogging in a postgraduate course in Mathematics Education. Makri and Kynigos integrated blogging as a medium of asynchronous communication and reflection impulse. The researchers focused on the study of the participants' role and practices' changes, using findings from dialogue and social practices analysis. Methodologically they adopted the Garrison & Anderson's (2003) framework for the analysis of the online learning systems, which considers the following decisive quality elements: the social, cognitive and teaching present. From the preliminary research data analysis the researchers notice, among others, the possibility of the development of a long-lasting "warm"-informal dialogue and non-monologue narration emerging collaboratively. According to the researchers the discussion genres expose structured cognitive present and constitute a significant evidence for the learning value of blogging. A straightforward comparison of the documented blogging case to the online forum technology could further clarify the selection criteria between the two technologies that are not noticeably contradistinguished in the above mentioned case study.

In the direction of educational exploitation of social software this paper presents a group blogging case study concerning seven students collaborating for the design of technologically enhanced Geometry learning activities for young children. The choice of blogging technology in this case is based on the claim that, to a large extent, it fulfils the communication and information management requirements of the learning by design pedagogical approach. The importance of the adoption of a well-defined pedagogical approach for the successful integration of blogs and any other content management model (e.g. wikis) is mentioned also in Chen et al. (2005) where a case of blogging integrated with the learning portfolio approach is described.

The main aim of the authors is to test the hypothesis that educational blogging in combination with the appropriate pedagogical approach could enhance the quality of learning and teaching at least for Mathematics education in tertiary level. Goals of the implemented intervention are: investigation of students' familiarisation ability to blogging, assessment of the "learning by design" approach compatibility to blogging technology, investigation of the quality of the obtained learning and the analysis of the interaction developed by the participants.

In the following the theoretical framework for the case study is presented initially, then the research conditions and the main research questions are reported, next the research evidence are presented along with authors' interpretations and finally a general discussion is provided.

## **Theoretical framework**

### **Learning and language**

Our basic theoretical framework stems from the sociocultural view of learning; this theory (or to be more precise this group of theories) is mainly based on the work of Vygotsky (1986) and stresses the importance of social interactions in the establishment of knowledge. Language plays a basic part in this process, because it is more than a means to express and communicate the persons' ideas; it is the means by which they think and learn together (Mercer, 1995). Written speech in particular, given its systematic organization and intentionality (Duval, 1999), can be a powerful means to examine not only cognitive, but also communicative functions.

The sociocultural approach is related to collaborative learning, which in turn involves participants' interactions. An important aspect of all human interactions is, according to Goffman (1972), the participants' face, i.e. the positive social value that one claims for him/herself. In simple words, when people interact, they firstly consider how they will protect themselves from possible verbal (or non-verbal) threats. Most of the times people also consider the other participants' face, so they implement various politeness strategies (Brown & Levinson, 1987) in order not to pose any threat to them. The above remarks need to be taken into account during the analysis of the participants' roles even in the blog medium.

The design of the experimental part of the study is based on the hypothesis that educational exploitation of blogging should better be combined with an appropriate pedagogical approach. This combination leads to a controllable goal setting and educational efficiency assessment. In this study, the blogs seen as a communication medium are considered compatible to "learning by design" pedagogical approach (Orey, 2001) and to the development of a special writing, dialogue and reflection community.

## **Learning by design**

Learning by design is related to constructionism, according to which new knowledge is more effectively developed by students when they are actively engaged in the construction of an external, shareable artefact that helps them to reflect and collaborate. Learning by design emphasises the learning value of the artefact design and at the same time underlines the learning benefits of the process regardless of the final product. In learning by design activities the designed artefacts are of personal significance for the students and 'represent' the learning outcome. There are several views on what constitutes learning by design; according to Han & Bhattacharya (2001) learning by design environments include: authenticity of the design theme, a balanced mixture of constrained, guided, scaffolded challenges and open design tasks, a rich variety of feedback information for designers, discussion and collaboration, experimentation, inquiry and reflection.

During learning by design activities students follow repetitively, in general, the following stages: theme-question selection; target group description; design and first implementation of the artefact; pilot application of the artefact; feedback information gathering; reflection; and design adaptation. Learning by design activities foster interaction between participants (students and teachers). The teacher's role includes supporting the students' efforts and guiding their interaction in order to obtain the desirable learning result. Teacher needs to supervise systematically the group interactions and decide for the quality of their verbal exchanges. In a conventional classroom students' interaction is often not allowed at all (frontal teaching); furthermore, in the case of non online group-teaching, teachers can not easily monitor and intervene during the collaboration, focusing mainly on the output-product of the activity. The employment of an online environment seems to be an efficient solution for managing the increased number of learning interactions occurring in group collaboration and for the improvement of the provided education quality. The selection of the appropriate online environment is considered to be an open problem; in the following, the specific features of the blogs that ground their selection for the present case study are described.

## **The blogs**

The weblogs or just blogs are World Wide Web (www) sites where a group of users can publish hypermedia articles (posts), presented in reverse chronological order. Modern blogging systems provide automatic chronological archive of posts, search service based on key words, user defined tags annotating posts, tag based indexing of articles etc. As far as the interactions facilities are concerned, blogs permit users and readers to publish comments on articles. Furthermore, blogs provide mechanisms for inter-blogging communication, such as distribution and re-publication of articles (rss syndication), and mechanisms for tracking the references to specific articles in other blogs (TrackBacks). In other words, it is possible for a blog to automatically receive information for the references to its articles and publish digests of the corresponding articles. All this interchange between blogs interconnects them in a hyper structure known as blogosphere.

Blogs as described above constitute a content management and an online publishing model capable of fulfilling the requirements of many different applications ranging from simple internet portals, newspapers, newsletters, to personal web sites. Additionally, the blogs' minimum requirements for technological skills make them quite popular. Many technological solutions for the implementation of blogs are available, including open source or proprietary stand alone systems, free blogging services available mainly from major internet search engines and blogging solutions embedded in learning management systems like moodle. All this spectrum of solutions practically means that internet access is the only essential prerequisite for educational blogging.

## **Educational features of blogs**

What are the educational uses of blogs? How can a teacher integrate them to a real course? In an educational technology wiki hosted at <http://edutechwiki.unige.ch/> it is suggested that students can use blogs in order to:

- take notes from the classes
- collect learning resources and share ideas and experiences
- log notes and observations during an inquiry learning activity
- manage a project
- publish news and information about the course, like course syllabus, calendar, handouts etc.
- develop dialogue like in an online forum

- reflect and communicate with teachers and peers-students
- improve their writing skills
- develop collaboration and social skills
- obtain the motivation of writing for readers who comment you in order to participate more actively in the course; and
- run online school newspapers

Teachers can utilise blogs in order to increase the communication among the participants of the course as well as the level of their participation and the depth of engagement. Blogs support students' collaboration and enable teachers to monitor the evolution of students' interactions, intervene whenever needed and obtain diagnostic information during the implementation of the assignments. Additionally, the teacher can store selected learning resources to a class blog in order to express and form the desirable direction.

As far as the support of learning by design activities is concerned blogs seem to fulfil the requirements of every distinct stage of it. More specifically, in the stages of theme selection and target group definition students can use blogs as a dialogue medium giving equal opportunities for expression to all. Students can also document target group characteristics producing hypermedia articles and also improve them by commenting and reviewing. In the stages of design, pilot application and revisions, blogs can serve as the means for the revision process by producing several document versions through the automatic archive and the tag and keyword-based search facility. An important feature at every stage is the projection of the contribution of each participant separately. The comparison of communication and information management requirements of learning by design activities to blog facilities clarifies the advantages of blogging for the support of such kind of activities especially to non-technologically supported implementation. The selection of blogs instead of other social software applications for the support of learning by design can be justified comparing their basic features:

**Blogs vs. simple html pages:** The result of this comparison is rather obvious. In simple html pages there is no automatic chronological archive service, no articles indexing according to tags, no separation of each participant's contribution etc. The simulation of these services demands advanced technological skills prohibitive for most students.

**Blogs vs. online forums:** Forums have a narrower application range than blogs. Forums' main advantage is the graphical representation of the time and of the participants' relationships in a dialogue using a tree structure. Forums are not designed for the posting of extensive articles; unlike blogs they usually do not permit the publishing of simple pages, link directories and resource collections. Furthermore, blogs can exploit plugins in order to provide more specific services for education like geographical information management, photo albums, resource centres etc.

**Blogs vs. wikis:** The basic idea behind the wiki content management model is the collaborative authoring of an online document. Despite the page version comparison and log facilities there is no automatic chronological archiving of the articles, making it more difficult to support the repetitive stages of learning by design activities.

**Blogs vs. Content Management Systems:** General purpose CMS demand unnecessary complexity and familiarisation load for the typical educational use.

**Blogs vs. Learning Management Systems:** Blogs and LMS are in a rather complementary relation, because blogs can be used as building blocks for instructional design in LMS. Embedding blogs in LMS gives the additional advantage of controlling the readers' population and avoiding side effects like spam email. On the other hand, LMS reproduce the traditional educational relationships and are not appropriate when the goal is to transform the relations and provide conditions for more voluntary and informal engagement of the participants.

## Research

### Conditions

The participants in the research were the first author as teacher and seven volunteer students from the audience of a University course called "ICT applications and product development for mathematics instruction". In the following, students are referred using nicknames. Students' participation in the blog fulfilled their project obligation for the

above course. The particular course aimed to enable students a) to get familiar to ICT applications in young children mathematics education, and b) to design and develop simple educational applications and learning activities for technologically enhanced mathematics learning environments. After having been informed about the research the students participated in a three-hour training for blog use. All the communication and interaction among the students had to be accomplished through the blog since there were no face-to-face meetings of the group.

The whole project that was assigned to the students was divided in the following four phases:

**Phase 1.** Students study some articles (Clements & Sarama, 2000a, 2000b, 2002; Fessakis & Tassoula, in press) about geometry education for young children, and the use of ICT for mathematics education. At the end of phase 1 the students had to post an article to the blog summarizing the papers they studied along with their own thoughts and comments. This phase aimed to set a more or less common theoretical view about the opportunities of ICT exploitation for young children geometry learning.

**Phase 2.** Students familiarize themselves with the use of the “Ladybug Leaf” and “Turtle Geometry” java applets from National Library of Virtual Manipulatives at Utah State University (NLVM, 2007) inspired from logo programming language and developmentally adapted for young children. At the end of this phase students had to post a review article to the blog putting together their thoughts about the applets (advantages, disadvantages, opportunities for utilisation, difficulties, developmental appropriateness etc.). Students were encouraged to comment on each other’s articles.

**Phase 3.** The third phase is the most demanding in terms of mental and collaboration effort required. Students had to design common classroom learning activities making use of the given java. In addition, students had to peer review and comment on the designed activities proposing corrections, improvements extensions etc. The goal of the group was to publish a common set of activities as their collaborative final product.

**Phase 4.** In the last phase students assess their collaboration and participation in the blog in general by posting a final informal article.

The duration of each phase was about one week except from phase 3 that lasted two weeks because of its work load and importance. The above learning by design activity was implemented during the second semester of the academic year 2006-2007.

### **Technological platform**

The course blog was implemented using a free blogging service of a Greek portal which can be found at URL: <http://gfsakis.pblogs.gr>. The selection of the specific platform is based mainly on the availability of a user friendly article and html pages editor with a satisfactory Greek language support. The platform was adequate for the purposes of the research, but for general educational use blogs embedded in LMS are preferable.

### **Research hypotheses and questions**

The basic hypotheses for the design of the above described intervention are:

1. Students can rather easily get familiarised to the use of blogs in such a degree that they can successfully participate in corresponding learning activities.
2. Blogs can efficiently support the communication and information management requirements of learning by design activities in such a degree that participating students can have increased opportunities for: a) meeting high quality learning experiences; b) communicate and interact with their peers and teachers; c) get feedback information in order to review and revise their designs; and d) develop a dialogue relative to the content.



## Methodology

The presented research constitutes basically a case study. The research approach is based on the design-based research model (Collective, T. D.-B. R., 2003) and the view that mathematics education can be considered as a “design science” (Wittmann, 1995). According to Wittmann (1995) mathematics education is more like engineering because they both study artificial objects in contrast to e.g. Physics which studies natural environment objects. As a consequence, Wittmann proposes that significant mathematics education research results have the form of well-designed and empirically studied instruction units, consistent to fundamental education theoretical principles.

### Research data collection instruments

The research data collected and analysed include the user-to-blog interaction log files, the published content of the blog (articles and comments), the questionnaire that students answered upon the completion of the activity and an audio-recorded students’ group interview by the teacher and the second author that was also organized after the completion of the activity.

### Research data analysis-findings

The research data has been analysed according to the following four axes:

**Axis 1.** Learning outcome: The learning outcome assessment of the students’ participation to the specific learning by design activity is undertaken through the produced learning activities and the quality criteria proposed by the students.

**Axis 2.** Blog content analysis: The content of the blog, i.e. the participants’ contributions are analysed concerning their content and their communicative and interpersonal purpose.

**Axis 3.** Participants’ views: Upon the completion of activity students answered a questionnaire and participated in a common round-table conversation. Students’ view about their experience is documented analysing the corresponding research data.

**Axis 4.** Quantitative summary of participants’ interaction: Interaction log files of blogs usually contain data like the time and duration of user connections, the number of articles and the corresponding comments, etc. Processing these raw data makes it possible to gather useful information about the engagement of each student, the level of interaction etc. This information gives an overview to the teacher and can form a kind of mirror for the group.

In the following the main findings for each analysis axis are presented.

#### Axis 1. Learning outcome

In learning by design the learning outcome is mainly expressed by the final design. The detailed presentation of the students’ designed activities is out of the purpose of this paper. The learning outcome is going to be approached by the students’ answers to questions asking them to present the best and worse learning activities according to their opinion. We adopted this approach because it shows that students did not only manage to design interesting and high quality learning activities but in addition they became more conscious of the evaluation criteria that could be used for the assessment of such activities. Students invoked several assessment criteria like: to be interdisciplinary (e.g. combining arts and mathematics), to be related to everyday authentic themes, to create a pleasant atmosphere, to concern mathematical concepts [e.g. counting and shape recognition], to be clear, conceivable for children and feasible with the specific technological tools. The most popular learning activity, according to students’ answers, was the use of applets for the production of figures for the composition of a collage to decorate the classroom. Students’ answers to the question about the worse proposed learning activity refer to six different learning activities. Some characteristic answers are the following:

**A12E15. Which Geometry learning activity from those proposed in the blog did you like the most and why?**

**Valia:** *I mostly like the combinational ones, more specifically the activities which concern orientation, counting [and] geometrical shapes recognition. I find more appropriate those that are based on maze escape.*

**Tassos:** *I liked one of John's entitled "painting construction using geometrical shapes" because except from the mathematical learning goals it combines arts education.*

**A12E16. Which Geometry learning activity from those proposed in the blog did you least like and why?**

**John:** *The activity I least like is one of Kate's about perspective because I didn't manage to understand its main idea and purpose. I also, cannot understand how it is going to use the java applets since they do not provide the corresponding functions.*

**Evi:** *The activity I like the least is one of Valia's entitled "Paths" because I think it is complicated.*

**Mina:** *The "Complete the shape" [one of Roula's] because whatever the child configures is going to be a shape but not necessarily a triangle, square etc., which is the real intention of the activity designer. The instructions for the children have to be studied carefully.*

**Axis 2. Blog content analysis**

The data analysis was performed in two stages; initially, every article and every comment was categorised according to its content and its communicative purpose. Regarding the comments, we decided to categorise only the comments that had an informative purpose (and not for example the few ones that had a personal character). We ended up with the following categories:

**INF1:** Original article

**INF2:** Enrichment of one's own original article

**INF3:** Comment on someone else's article

**INF4:** Reply to a comment on one's own article

**INF5:** Reply to a comment on someone else's article

Table 1 shows each participant's post in the blog, concerning the above categories:

*Table 1. Number of publications per category*

	INF1	INF2	INF3	INF4	INF5
<b>Mina</b>	4	0	21	2	4
<b>Tassos</b>	5	2	6	2	1
<b>Valia</b>	3	1	13	0	0
<b>Evi</b>	5	0	3	2	1
<b>Roula</b>	4	0	0	0	0
<b>Kate</b>	4	5	10	4	1
<b>John</b>	3	0	7	0	0

It is obvious from Table 1 that there has been a differentiation between the participants concerning the number and the type of their posts. The INF1 category contains at least three articles per participant, which were defined as the minimum needed. Only three students enriched their articles (INF2), with Kate being the most active in this category (and in general). The category INF3 contains a big number of posts, since all students – with the exception of Roula – posted many comments on each other's articles. The posts of the INF4 category are relatively few, since there were comments that didn't receive an answer. The students justified this by saying that they found it difficult to be informed about recent comments in non-recent articles. The posts of the category INF5 are even less; this demonstrates the limited participation in this kind of exchanging comments.

Figure 1 demonstrates each student's participation according to these categories.

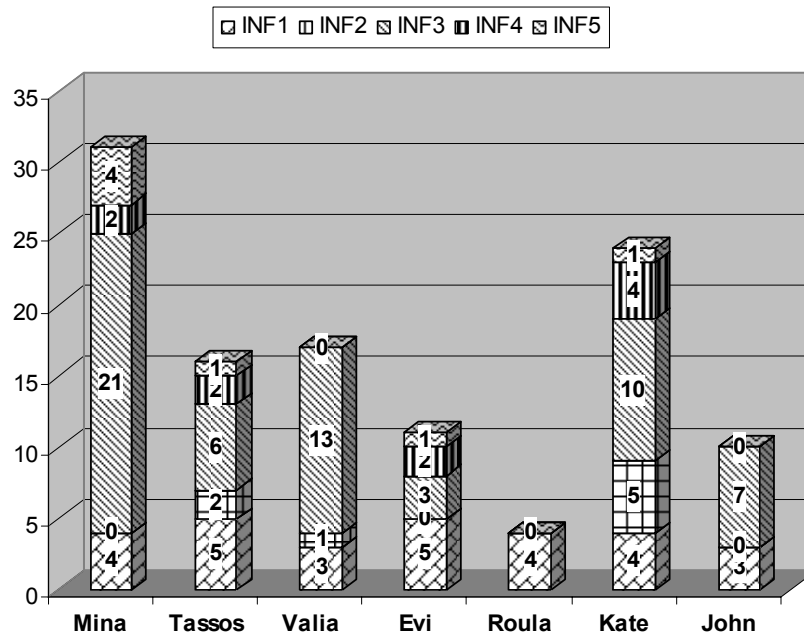


Figure 1. Students' contribution per post/comment category

In the second stage of analysis we adapted Mercer's (1995) methodology in order to study each contribution's function in two levels. In the first level, we categorised each comment into educational, academic and everyday discourse. In the second level, we examined the speech acts that were performed by each comment, mainly bearing in mind the face-saving perspective.

Concerning the first level, there were many instances when the students used academic discourse:

"Learning aims: Orientation, understanding of the concept of degree and its correspondence to the relative motion." (Evi, 29/5)

The above excerpt contains a list of concepts expressed in an impersonal manner without using any verb. Many participants (including the teacher) also used educational discourse in their posts:

"Kate, your first activity is very good. I believe that it can be easily comprehensible by an infant... Moreover, in this activity children can make comparisons and get familiarised with the notions of "smaller" and "bigger"." (Valia, 14/6)

Apart from an initial positive evaluation, the above excerpt contains a summary of the whole idea; this is a common strategy used by teachers when they evaluate their students' views. Everyday discourse was rarely found in the informative posts and this is an indication of the students' attitude towards the educational use of the blog.

The second level of analysis revealed the students' communicative and interpersonal intentions. In other words, we have found rich evidence of the students' concern to respect their fellow-students' face. The vast majority of comments – even the harshest ones – were based on the scheme: initial positive remark followed by the basic remark:

"John, your ideas are interesting... *I would suggest not to use the term collage since you are not going to stick anything. Moreover, the term "knowledge: does not seem appropriate; it's better to use the terms familiarisation, informing, briefing when you talk about the aims ...*" (Mina, 29/5)

The above excerpt, apart from being based on the scheme that we mentioned, it also contains what Brown and Levinson (1987) call politeness strategies; the words marked in italics all serve the same purpose: to “soften” the threat to the reader’s face produced by the information that follows (see e.g. Tatsis & Rowland, 2006).

Another remark, which connects the two levels of our analysis, is that we believe that the participants incorporated the academic and the educational discourse in their posts in order to maintain or even strengthen their own face. Academic discourse has a high status in the educational community, thus it can be seen as a tool to strengthen or maintain one’s position. Another strategy towards this direction was the avoidance of answering to comments or to the teacher’s prompts for further investigation on particular issues.

### Axis 3. Participants’ views

The participants’ views about their blogging experience have been recorded using the final questionnaire and the common round-table conversation. The students’ views analysis is organised according to the following four dimensions:

- 3a. writing for readers-commentators and peer reviewing
- 3b. the features of constructive comments
- 3c. appropriateness of blogs for the pedagogical approach
- 3d. general opinion and difficulties

Open questions with unfinished sentences have been used for the examination of the students’ views. In the following the most significant findings are presented per sub axis.

#### 3a. Writing for readers-commentators and peer reviewing

From answers to questions:

A3E1. Writing articles for my fellow-students as readers...;

A3E3. Writing comments for my fellow-students articles..., and

A3E5. The comments of my fellow-students for me ...,

it seems that students were initially hesitant about posting articles and especially about commenting their fellow-students’ articles. They state that initially they were uncomfortable and seriously thought about the possibility of being misunderstood. Finally, most of the students overcame their diffidence; in addition some of them state that they were excited by the exchange of opinions. Despite the general enjoyment and excitement one student did not post any comment and another declared that he posted intentionally neutral comments in order not to cause bad feelings and conflicts.

Students in general, believe that the comments they received helped them to identify their mistakes and improve their designs. In the negative features of the comments students stated that they were not always informative and some times others seemed to hesitate in expression of judgments. Some students declared that they could produce more comments if they had internet access from their homes and/or there were more participants to the blog.

Some characteristic answers are quoted:

#### A3E2. Writing comments for my fellow-students’ articles ...

**Valia:** *was not an easy task because it was only based on our personal judgement rather than on a general model.*

**John:** *I can say that they contributed in the improvement of my opinion about some questions....*

**Kate:** *It was difficult especially in the beginning because there was the risk of misunderstanding. I was careful about the way of expressing my opinion.*

**Mina:** *It made me feel uncomfortable initially but in the course I liked the process of commenting on mine and my fellow-students’ articles.*

**Tassos:** *It puzzled me in the beginning because I did not like to comment negatively my colleagues despite the fact that it could help them improve their errors.*

**Roula:** *It was something that I did not do many times over [actually she did not post any comment].*

### 3b. The features of constructive comments

An interesting point is the students' views about the features of constructive comments. To investigate this issue, students answered the completion question A3E10. The comments help more when... Students mention that comments should: be well-intentioned, have friendly style, be clear, mark problems while proposing solutions, aim to continuous improvement. Some characteristic answers are quoted:

#### A3E10. The comments help more when ...

**Valia:** *They are clear and potentially propose some ways to face a problem.*

**John:** *they aim to get you improved by vital remarks. Even for the best article you should prompt the author for improvement.*

**Evi:** *They are honest despite of the interpersonal relations of the participants.*

**Kate:** *They are guiding and mention the mistakes.*

**Mina:** *They are well-intentioned, informative and written by people having knowledge about the article theme.*

**Roula:** *They have a friendly-consulting style.*

**Tassos:** *They point out mistakes we have done in some activities helping us to improve them.*

### 3c. Appropriateness of blogs for the pedagogical approach

The students' point of view about the appropriateness of blogs publishing model for the support of learning by design activities is investigated through the answers to the question A4E6. If the blog was not available for the implementation of the same work then... Students' answers show that they understand the communication requirements of the pedagogical approach as well as the value of the capability of monitoring the whole process during the activity. Students point out the difficulties they could have in order to organise the required face-to-face meetings; in addition students mention that without the blogging system they could not comment each others' articles in such an extent and depth while they were continuously monitoring the evolution of the proposed designs

The quoted answers of students are indicative:

#### A4E6. If the blog was not available for the implementation of the same work then...

**Valia:** *probably we could not have the feeling of collaboration and responsibility to each other so intensively.*

**John:** *in order to manage taking our collaborators' opinions we should arrange meetings for all the members. Even if this was feasible we were going to face difficulties because of the lack of the capability to continuously monitor the work of others.*

**Evi:** *we could not easily have our own designed activities commented and articles by the rest of the participants.*

**Kate:** *we would face a coordination problem. Each one of us has many obligations and blogging is highly demanding in terms of the required meeting.*

**Mina:** *we were going to need much time, many transfers or the use of some other software that would demand special skills.*

**Roula:** *we were probably going to use the "traditional" methods and lack this experience.*

**Tassos:** *we wouldn't have the chance of direct communication between us. (Comments between us)*

### 3d. General opinion and difficulties

In questions about their overview of the blogging experience students were satisfied in general with their participation, most of them mention that they will participate in group blogging in the future while some of them intend to run a personal blog as well. Students' proposals, for the improvement of similar activities in the future, include extending of the duration, increasing of the number of participants, and developing a new activities notification service. The lack of a notification service was a disadvantage of the selected technological platform. During the group interview students mentioned that it was difficult for them to know about new comments in their "old" articles or comments. Students invoked this awareness difficulty as a major reason for some unanswered comments and questions that were observed in the blog.

#### Axis 4. Quantitative summarization of participants' interaction

The analysis of research data in this axis aims at quantitative and graphical summarisation of participants' interaction. The following graphical representations of the interaction raw data stored in the system log file gives an overview of the group work course at a glance. Two diagrammatic representations are presented: the first is a bubble chart summarising the article posts while the second is a social network chart depicting the commenting relations of the participants. During the learning by design activity the participants posted over 50 articles and 100 comments in total.

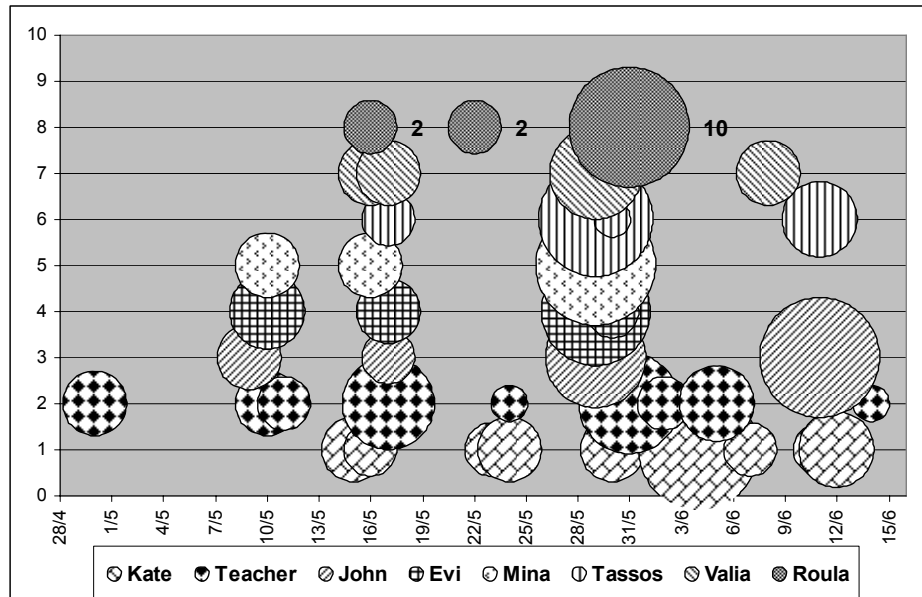


Figure 2. Articles production synopsis

The bubble chart of figure 2 constitutes a projection of the blog to the articles posting dimension. The horizontal axis corresponds to the publishing dates while the vertical to participants. Bubbles with their centre in the same height of the vertical axis have the same fill pattern and represent the article posting of a specific participant. Articles of a particular participant posted within the same day are depicted by one bubble the diameter of which is the sum of the diameters of the bubbles for each article. The numbers next to the bubbles refer to the number of the comments posted for each corresponding article increased by 1 (in order to have a bubble even for uncommented articles). In other words, the diameter of the bubbles is proportional to the impact of the articles in terms of comments. The height of each participant's line of bubbles is related to the total number of comments his/her articles received. More specifically, the lowest (heavier) line (Y=1, Kate) has received the most comments for her articles while the highest (lighter) line (Y=8, Roula) the least.

The bubble chart depicts the contribution of each participant to the blog in terms of articles and corresponding comments. It gives a kind of relative weight measure for the articles' impact. The specific chart shows a significant number of articles and an interaction intention increasing by time. This is depicted by the diameters of the bubbles which get longer by time. This is consistent with the students' statements that initially they hesitated to publish but felt comfortable in the course. It is also possible to see the deadlines of the project because of the accumulation of the bubbles in the corresponding dates (e.g. 17 and 29 of May).

The bubble chart gives an overview of the articles production, their impact and the participants' relative contribution for this but it contains no information about who is commenting whose articles, what is the participants' relative contribution on comments, etc. The comments dimension of blogging is equally important to the articles one because it includes the basic mechanism for feedback information supply and reflection prompt. The blog activity is projected to the comments dimension using the social network graph presented in Figure 3. The social network graph has been produced using the Ucinet v6 software system (Borgatti et al., 2002). The graph shows the comments production for the specific blog.

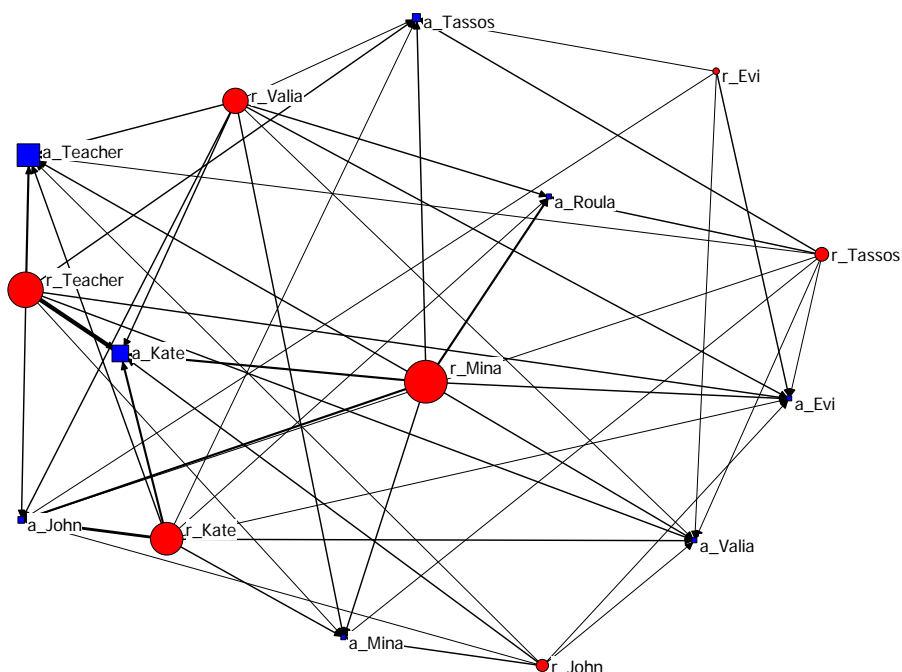


Figure 3. Comments production synopsis.

The nodes shaped as circles represent the participants as commentators-reviewers (labels with prefix “r\_”) while nodes shaped as squares represent the participants as article authors (labels with prefix “a\_”). The size of the nodes is proportional to the total number of comments and articles of the corresponding participant for the circles and squares respectively. The arcs are directed from the commentator to the corresponding author. The thickness of the arcs is proportional to the total number of comments for the corresponding couple of participants. It is worth commenting that in this version of the diagram the thickness is related only to the number of comments and contains no information about the quality of it (positive or negative).

By observing the graph it is obvious that *r\_Mina* posted the most comments of all, and *r\_Kate* takes the second place on comments production. It is also apparent that all participants with the exception of Roula have posted comments to every other. Roula did not post any comment since there is no circle node labelled *r\_Roula*, although her articles received comments. As far as the topological orientation of the nodes is concerned, the left part of the graph hosts in general the participants who were more productive in comments and the authors (squares) are located near their more frequent commentators. The optical observation of the graph gives rich information about the commenting interaction of the blog. The viewer can readily find out who is commenting whom and in what frequency. It is also possible to compare the relative contribution of the number of comments of each participant. In the specific graph we see a high degree of connectivity that is a strong evidence of the appropriateness of the learning by design approach which required peer review of the designs fostering the comment production.

Of course the simple quantitative overview offered by the above graph allows a useful estimation for the intensity and extension of the interaction but does not allow an accurate quality assessment of the participants’ learning experience; this assessment needs the combination of the content review and analysis.

## Discussion and Conclusions

In this paper a case study for the use of blogging in a learning by design project is described. The reported work is inspired by the use of social software for the construction of teachers’ professional development communities of practice (Marcelo & Bairral, 2007), while at the same time considers the results about social practices of teachers when blogging that are reported in Makri & Kynigos (2007). The presented case study extends the mentioned works focussing on the particular case of group blogging as a communication and information management system in combination with the specific pedagogical approach of learning by design. The main idea behind the study of

blogging use for collaborative learning by design was to investigate the education enhancing opportunities of such a constantly gaining in popularity software system.

As far as the familiarisation of students with blogs is concerned, the successful completion of the activity, the articles, and the comments' quantity and quality shows that despite the initial difficulties, students quickly acquired the required skills for blog use.

As far as the appropriateness of blogs for the implementation of learning by design activities is concerned, students had increased opportunities to receive feedback information in the form of fellow-students' and teachers' comments. Furthermore, the participants have described the features of constructive comments that helped them to overcome their initial hesitation about the peer review and commenting. The blog support for the learning by design activity communication and information management was decisive, according to the students. The blog enabled them to coordinate and communicate efficiently without the need to meet face-to-face.

As far as the learning outcome is concerned, in learning by design activities the learning outcome is mainly represented by the final design of the students. Students managed to design a variety of learning activities which were revised according to their fellow-students' comments. The successive, in general, understanding of the learning activities designed by others and the accurate critique, as it was recorded in students' answers to the last questionnaire shows a rather satisfactory learning outcome. The students' participation in the blog has given them the opportunity to experience a rich set of learning activities ideas in order to obtain the general learning activities quality assessment criteria declared in their questionnaire answers.

The content analysis has shown that all participants implemented some face-saving strategies in order to protect their own and their fellow-participants' face. This has not restrained them from posting meaningful and useful remarks that contributed to the ongoing learning activities design and discussion on it. For the teacher it proved easy to monitor these discussions and intervene whenever necessary.

The above remarks lead to the conclusion that the students have experienced a high quality learning experience, which contained a high density and long duration dialogue relative to the learning content, as well as opportunities for collaboration and reflection.

Improvements for similar future applications could include: increased number of participants, longer duration, and the development of a new blog activity notification service briefing any new comment or article for each participant separately. As for the research dimension, the presented work could be continued with more sophisticated interaction analysis techniques, the consideration of other pedagogical approaches, and the comparison of blogs to other social software.

Our conclusion is that blogs exploitation in combination to the proper pedagogical approaches has the potentiality to enrich the quality of future teachers' education.

## References

- Allen, C. (2004). *Tracing the evolution of social software*, retrieved June 15, 2008 from: [http://www.lifewithalacrity.com/2004/10/tracing\\_the\\_evo.html](http://www.lifewithalacrity.com/2004/10/tracing_the_evo.html).
- Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis*, Harvard, MA: Analytic Technologies.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*, Cambridge: Cambridge University Press.
- Chen, H. L., Cannon, D. M., Gabrio, J., & Leifer, L. (2005, June). Using Wikis and Weblogs to Support Reflective Learning in an Introductory Engineering Design Course. *Paper presented at the 2005 American Society for Engineering Education Annual Conference & Exposition, Portland, Oregon*, retrieved June 15, 2008 from: [http://rice.stevens.edu/fileadmin/rice/pdf/ASEE2005\\_Paper\\_Wikis\\_and\\_Weblogs.pdf](http://rice.stevens.edu/fileadmin/rice/pdf/ASEE2005_Paper_Wikis_and_Weblogs.pdf).



- Clements, D. H., & Sarama, J. (2000a). Young children's ideas about Geometric shapes. *Teaching Children Mathematics*, 6 (8), 482-488.
- Clements, D. H., & Sarama, J. (2000b). The earliest Geometry, Teaching Children Mathematics. *Teaching Children Mathematics*, 7 (2), 82-86.
- Clements, D. H., & Sarama, J. (2002). The role of technology in early childhood learning. *Teaching Children Mathematics*, 8 (6), 340-343.
- Collective, T. D.-B. R., (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32 (1), 5-8.
- Downes, S. (2004). Educational blogging. *EDUCAUSE Review*, September/October, 14-26.
- Duval, R. (1999). Ecriture, raisonnement et découverte de la démonstration en mathématiques. *Actes de la X-ème Ecole d'Eté de Didactique des Mathématiques, II*. IUFM de Caen, 29-50.
- Ernest, P. (1994). What is social constructivism in the psychology of mathematics education? In. da Ponte, J. P & Matos, J. F. (Eds.), *Proceedings of the 18th Annual Conference of the International Group for the Psychology of Mathematics Education*. Lisbon, Portugal: University of Lisbon, Vol. 2, 304-311.
- Fessakis, G. & Tassoula, E. (in press). Design of an educational robot remote controlled by young children via computer software for the development of mathematical concepts and space perception skills. "*Astrolavos*" *Journal of Greek Mathematical Society*.
- Garrison, D. R. & Anderson, T. (2003). *E-learning in the 21st century: A framework for research and practice*, London: Routledge Falmer.
- Goffman, E. (1972). *Interaction Ritual: Essays on Face-To-Face Behaviour*. Harmondsworth, Middlesex: Penguin.
- Han, S., & Bhattacharya, K. (2001). Constructionism, Learning by Design, and Project-based Learning. In Orey, M. (Ed.), *Emerging perspectives on learning, teaching, and technology*, retrieved June 15, 2008 from: <http://www.coe.uga.edu/epltt/LearningbyDesign.htm>.
- Kim, B. (2001). Social constructivism. In Orey, M. (Ed.), *Emerging perspectives on learning, teaching, and technology*, retrieved June 15, 2008 from: <http://www.coe.uga.edu/epltt/SocialConstructivism.htm>.
- Makri, K., & Kynigos, C. (2007). The Role of Blogs In Studying The Discourse And Social Practices of Mathematics Teachers. *Educational Technology & Society*, 10 (1), 73-84.
- Marcelo A., & Bairral, M. A. (2007). Building a community of practice inquiry about geometry: A study case of pre-service teachers interacting online. *Interactive Educational Multimedia*, 14, 40-53.
- Mercer, N. (1995). *The Guided Construction of Knowledge: Talk amongst teachers and learners*, Clevedon: Multilingual Matters.
- NLVM Utah State University (n. d.). *National Library of Virtual Manipulatives*, retrieved June 15, 2008 from: <http://nlvm.usu.edu>.
- Orey, M. (2001). *Emerging perspectives on learning, teaching, and technology*, retrieved June 15, 2008 from: <http://www.coe.uga.edu/epltt/CognitiveApprenticeship.htm>.
- Papert, S. (1993). *The Children's machine: rethinking school in the age of the computer*, New York: Basic Books.
- Tatsis, K., & Rowland, T. (2006). Vague language in Greek and English mathematical talk: A variation study in face-work. In Novotná, J., Moraová, H., Krátká, M. & Stehliková, N. (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education*, Charles University, Vol. 5, 257-264.
- Vygotsky, L. S. (1986). *Thought and Language*, Cambridge, MA: The MIT Press.
- Wenger, E. (1998), *Communities of practice: Learning meaning and identity*. Cambridge, UK: Cambridge University Press.
- Wittmann, E. (1995). Mathematics education as a "design science". *Educational studies in Mathematics*, 29 (4), 355-374.

## A case study of enabling factors in the technology integration change process

**Pi-Sui Hsu**

Department of Educational Technology, Research, and Assessment, Northern Illinois University, 219B Gabel Hall,  
DeKalb, IL 60115, USA // Tel: +1 (815) 7536025 // phsu@niu.edu

**Priya Sharma**

Department of Learning and Performance Systems, The Pennsylvania State University, 314 C Keller Building,  
University Park, PA 16802, USA // Tel: +1 (814)8 654374 // psharma@psu.edu

### ABSTRACT

The purpose of this qualitative case study was to analyze enabling factors in the technology integration change process in a multi-section science methods course, SCIED 408 (pseudonym), from 1997 to 2003 at a large northeastern university in the United States. We used two major data collection methods, in-depth interviewing and document reviews. We interviewed seven participants who played the role of change agents twice during the six-year period of time and examined documents related to integrating science specific technology in this multi-section course. The formation of a sustained, shared leadership team, the formation of a learning community, and the positive influences from educational systems emerged as three enabling factors to facilitate and sustain the technology integration change process. The interaction of the three factors also appeared to encourage greater capacity to deal with change.

### Keywords

Teacher education, Educational change, Educational technology, Educational leadership

### Introduction

The influence of technology integration on broader educational change efforts is an important area of concern for the future progress of the field (Holloway, 1996; Surry & Ely, 2006). In light of the potential of technologies in helping students' understanding of new concepts, initial efforts focusing on technology integration in teacher education programs have been undertaken nationwide. From 1999, the Preparing Tomorrow's Teachers to use Technology (PT3) initiative has funded numerous teacher education programs in the United States. Technology integration into these field-based preservice teacher education programs have shown preliminary successes in modeling preservice teachers' use of technology, fostering their collaboration by use of technology, and increasing their confidence and enthusiasm in using technology (Brush, Glazewski, & Rutowski, 2003; Seels, Campbell, & Talsma, 2003).

However, the complexities of technology integration make it difficult to facilitate and sustain the change process. Although these reports shared successful aspects of the ongoing projects, they also indicated numerous barriers to integrate technology in teacher education programs. For example, faculty leaders encountered and overcame numerous difficulties, which included incompatible software purchases, limited access to computers, preservice teachers' and in-service teachers' initial resistance, and the lack of time (Brush, Glazewski, & Rutowski, 2003; Seels, Campbell, & Talsma, 2003). Furthermore, these reports did not address subject specific technology and were conducted in too wide of a context (program, general type of technology such as video production, HyperStudio, web sites) to provide knowledge about what strategies that faculty leaders could adopt to facilitate and sustain the process.

Thus, the purpose of this study was to analyze enabling factors in the technology integration change process in a multi-section elementary science methods course, SCIED 408 (pseudonym), from 1997 to 2003, the time this study was conducted, from the perspective of change agents. Havelock and Zlotolow (1995) defined change agents as "solution giver," "process helper," and "resource linker" (pp. 9-10). In this study, change agents are the faculty members and course instructors who initiated and led the technology integration change process.

We selected SCIED 408 for two reasons. First, in the last two decades, science education has embarked on a major reform, recognizing that technology integration in science education could enhance science teaching and learning (National Research Council, 2001). Research has addressed three major types of technology tools, referred to as inquiry empowering technologies or science specific tools (Linn, 2003; Windschitl, 2000; Zembal-Saul, Munford, & Friedrichsen, 2002) that are commonly used to assist elementary students in inquiring scientific understanding,

carrying out investigation, communicating, and developing products. Three types of technologies include data collection tools, simulations and modeling tools, and online collaborative tools. Data collection tools enable students to collect data efficiently and allow them to perform multiple analyses of the data set in a brief period of time. Simulations and modeling tools enable students to focus on exploration and representation of scientific processes and conceptions. Online collaborative tools are intended to create a learning community that students can share resources and explore scientific phenomena with communities beyond the confine of their own classrooms. Although these tools have shown potential to support students' development of science inquiry skills, they are rarely used or seen in college level science methods courses (Linn, 2003; Windschitl, 2000; Zembal-Saul, Munford, & Friedrichsen, 2002). SCIED 408 would be an exemplary case to study because the faculty members and course instructors made endeavors to integrate science specific technology tools into SCIED 408 from 1997 to 2003.

Second, Fullan (2007) suggested that “moderately complex changes of innovation integration can take from 3 to 5 years, while larger scale efforts can take 5 to 10 years” (p. 52). The six-year time period (1997-2003), suggested by Fullan, is a significant milestone to examine how the case had evolved. Thus, the elementary science methods course, SCIED 408, makes an appropriate case in this study. Examining this case could contribute to an understanding of the technology integration process in an educational context and lead to identification and understanding of the factors that support sustainable change. The major research question for this study was: What enabling factors might have helped to facilitate and sustain the technology integration change process from the perspective of change agents (faculty members and course instructors)?

## **Theoretical framework for inquiry**

We framed this research report by using systems theory as a basis to examine the factors and relationships that might have helped to sustain the technology integration change process in SCIED 408 from 1997 to 2003. The application of systems theory in education was pioneered by Banathy (1991, 2001) and popularized by Reigeluth (1994) as systemic change. Although a number of theoretical frameworks have discussed systemic change in different educational change contexts (Joseph, Jenlink, Reigeluth, Carr-Chellman, & Nelson, 2002; Rowland, 2004), empirical research is needed to examine the complexities, interrelationships, and interdependencies in a change process, particularly in the context of technology integration. Additionally, we identified three key components associated with systemic change—shared leadership, learning community, and educational systems—and we describe their contributions in the next section.

### **Shared Leadership**

Differing from a traditional view of leadership, systemic change recognizes that leadership is shared among several key persons in organizations rather than being dominated by one person (Dexter, 2005; Lambert, 2002; Mullan, 2002; Reigeluth, 1994; Rowland, 2004; Senge & Lannon-Kim, 1991). The leadership team can consist of key persons at different levels or departments, including faculty, teachers, principals, or administrators. In their study of the influence of leadership on reforming inquiry-based science education in 13 Chicago elementary schools, Spillane, Diamond, Walker, Halverson, & Jita (2001) indicated that the leadership team consisted of the principal, the science coordinator, and the assistant to the coordinator. The principal appointed a specific classroom teacher as a science coordinator. The science coordinator, who was funded by the school district, worked with an assistant provided by the principal and worked with science teachers across all grade levels to integrate more inquiry-based science instruction. To empower change in science education, the principal doubled the weekly science instruction period, secured financial resources, established connections to external resources, including local universities, colleges, science institutions, and external science consultants. Another example of a leadership team is identified in another elementary school in North Carolina (Nesbit, DiBiase, Miller, & Wallace, 2001). In both studies, the leadership team showed a division in the functions of design, administration, and liaisons, although some of the functions may have overlapped.

### **Learning community**

In systemic change, a learning community serves as a powerful incentive to sustain the technology integration change process (Joseph, Jenlink, Reigeluth, Carr-Chellman, and Nelson, 2002). According to Cibulka and Nakayama

(2001), a learning community is “a group of educators committed to work together collaboratively as learners to improve achievement for all students in a school...one that consciously managed learning processes through an inquiry-driven orientation among its members” (p.4).

A learning community learns and grows by collaborating on a variety of activities, including study groups, reflective dialogue, and research. In their qualitative study of an innovative change process in a secondary school in Queensland in Australia, Andrews and Lewis (2002) indicated that shared understandings that emerged from a professional learning community had an important impact on other teachers in a school change process. The professional learning community consisted of ten volunteer teachers, both junior and senior. The composition represented a wide range of backgrounds, experiences, and beliefs. Although the change process was full of ambiguity, the volunteers used shared dialogue across different departments to reconcile different perspectives on pedagogy, assessment, and ways to motivate students. In shared dialogue, they could think, reflect, draw on their professional experience, and listen to others’ ideology and belief systems. They also felt empowered to perceive themselves as guides for changing other teachers in the school. Each community member mentored two teachers and guided them through the change process. Another example of a learning community is identified in the FIRST (Fund for the Improvement and Reform of Schools and Teaching) project initiated in North Carolina (Nesbit, DiBiase, Miller, & Wallace, 2001).

### **Educational Systems**

In an educational context, different levels of educational systems in a systemic change process include the local community level, the university level, the state level, and the national level (Reigeluth, 1994). At the local community and university level, it is essential to involve local school district teachers and students by involving the local school community in university-held events and delegating liaisons to regularly coordinate activities involving community and school district personnel (Grove, Studler, and Odell, 2004; Taylor and Wochenske, 2001). At the state level, the department of education plays a critical role in the technology integration change process because it dictates curriculum design and funding of resources to sustain the changes (Brown, 2007).

At the national level, various activities are needed to sustain change, including gaining recognition from accreditation agencies; attending national contests, reviews, and meetings of professional organizations; disseminating experience by authoring articles in journals; and developing partnerships with educational software companies for appropriate support. Brooks-Young (2007) suggested the establishment of partnerships between the educational programs and vendors of educational software to ensure that students receive appropriate exposure to software. Brooks-Young further suggested the National Council for Accreditation of Teacher Education (NCATE) and other organizations that establish technology standards for teacher education standards could take the initiative to implement such arrangements with educational vendors.

The technology integration change process can involve different levels of systems. It is essential to identify the impacts of different systems on technology integration and to exercise strategies to reinforce methods to gain support from and tackle possible resistance from these systems.

### **Methodology**

This study used a qualitative research methodology, specifically case study research design, to explore answers to the research question. The first author was the major researcher responsible for sampling the participants, collecting the data and analyzing the data. The first author was involved in the teacher education program as a field experience supervisor and course instructor but was not involved in the technology integration change process in SCIED 408 directly. The first author never taught SCIED 408. The first author never worked with any participants in any fashion.

### **Context- Restructure of SCIED 408**

This study was conducted at a comprehensive, land-grant research university. The science education program is one of the programs in the Department of Curriculum and Instruction of the College of Education. SCIED 408, *Teaching Science in the Elementary School*, is a required course before Elementary and Kindergarten Education majors can

engage in student teaching. The course is taken concurrently with a mathematics methods course, a social studies methods course, and an elementary education field experience course. On average, six to seven sections of SCIED 408 are offered each semester. Approximately thirty preservice elementary teachers are enrolled in each section.

Based on a three-phase conceptual change model, SCIED 408 was restructured with three types of science specific technology tools (see Table 1). Three strands of science specific technology tools, including data collection tools such as PASCO Science Workshop probeware (PASCO Scientific, 2008) and Vernier PASPORT (Vernier Software Technology, 2008), simulation and modeling tools such as Tom Snyder Science Court (Tom Snyder Productions, 2008), and online collaborative tools such as Connecting Community of Learners (CCL) and A New Global Environment for Learning (ANGEL), were used in three phases in SCIED 408.

In the first phase of SCIED 408, preservice teachers experienced technology-rich science as learners and became familiarized with the aforementioned technology tools. In the second phase, preservice teachers taught a small group of children using exemplary science curriculum infused with appropriate technology tools under the guidance of the methods instructor in an event, Science and Technology Experience at P S (STEPS) Days, where the science education program invited local elementary students to the campus. In the third phase, in their field experience classrooms, preservice teachers taught technology-enhanced science lessons using technology resources from the program or from their field experience site. Throughout the course, preservice teachers reflected on their teaching by authoring web-based portfolios.

### **Data Collection**

There are two major data collection methods in this study: in-depth interviews and review of documents. This study used the criterion sampling strategy suggested by Miles and Huberman (1994) to select the people to interview. The purpose of the criterion sampling strategy was to select persons who had played or were currently playing the role of change agents. The following section provides a brief outline of the profiles of the seven participants and their qualifications and involvement with the project.

#### *Participant Profiles*

Dr. Zimmerman (pseudonym) joined the science education program as an assistant professor in May of 1997, the first year in the change process. She played a variety of roles in the change process of technology integration from 1997 to 2003: lead researcher, course instructor, course coordinator and consultant. As an associate professor, Dr. Donahue (pseudonym) had been in the science education program for a number of years prior to 1997 and played a number of roles: course instructor and the coordinator of the science teacher education program.

Dr. Stoker (pseudonym) joined the science education program on the state level grant in 1999, the third year in the change process, as liaison between SCIED 408 and the local elementary school districts. She was the course coordinator for SCIED 408 and had been a course instructor since 1999.

Ms. Hess (pseudonym) had been a course instructor from 1998 to 2001. In 1998, she conceptualized different technology projects with Dr. Zimmerman, planned how to integrate the technology projects into the design of the course, and trained the preservice teachers to learn these tools. Ms. Friedman (pseudonym) joined the science education program in 1998 as a doctoral student. Her major role in the change process was to co-author the state level grant proposal for technology integration in 1998 with Drs. Zimmerman and Donahue.

Ms. Aberdeen (pseudonym) joined the program in 1999 to pursue her master's degree and then Ph.D. degree. Between 1999 and 2001, she had been very active in the science education program, such as managing hardware and software in the science education program's computer laboratory. She had been a course instructor since 2001. Mr. Bell (pseudonym) joined the science education program in 2001 and had been a course instructor since that time.

#### *In-depth interviewing*

The first author conducted two rounds of interviews at different points of time (2001 summer, 2003 spring) in the technology integration change process in order to have a better understanding of the evolution of technology

integration in SCIED 408. The first author designed the interview protocols to indicate interview questions and closing comments (see Appendix A). The interview took about 90 minutes each. Before the interview, the first author presented each participant a copy of the informed consent form. The first author used a digital recorder to record the interviews and transcribed these interviews.

### *Review of documents*

Documents were the second source of data collection in this study. Merriam (1998) defined documents as “an umbrella term to refer to a wide range of written, visual, and physical material relevant to the study at hand” (p. 112). The main documents that were reviewed were the publications, the presentation papers, the technology integration state level grant proposal, the syllabi, and the class Internet resources. The publications and presentations provided information about each participant’s research interest, how these participants collaborated on research studies, and theoretical framework of technology integration in SCIED 408. The technology integration state level grant proposal provided detail on the design differences of SCIED 408 before and after the grant was awarded. The syllabi provided information regarding the assignments and activities related to technology integration in class. In addition, the Internet resources offered the names of the school districts where the partnerships were established.

### **Data analysis**

Data analysis followed the guidelines suggested by Miles and Huberman (1994), utilizing different levels of coding schemes. First, we coded transcripts with a start list of 24 codes for time (i.e., fall semester of 1999 as FA99) and technology integration project (i.e., simulations and modeling tools as SM), along with time and roles of people (i.e., SCIED 408 course instructor as CI, CC as course coordinator). The purpose is to describe the chronological history of technology integration and roles of people, which is a time series analysis strategy (Yin, 2002). The evolution of technology integration from 1997 to 2003 is presented in Appendix B. A matrix that details the roles of each interviewee during different periods of time in the change process is presented in Appendix C.

Second, we examined the matrices (see Appendix B and Appendix C) for trends and patterns of technology integration from 1997 to 2003. Following these trends, we identified interrelationships among faculty members, course instructors, and the technology integration projects in the change process by coding for themes, causes/explanations, relationships among people, or emerging constructs (Miles & Huberman, 1994). Thus, we generated a list of pattern codes. For example, three faculty members appeared to play a critical role consistently in the change process and thus were assigned a code.

Third, we then reduced, summarized, and synthesized, and integrated the data into an exploratory framework. In this step, we formed overarching themes by linking technology integration initiatives, faculty member and course instructors that emerged from the data analysis. For example, Appendix B shows that since 1997, a number of technology initiatives had been tried in class. By looking at enabling factors that sustained the change process, the data analysis indicated that three faculty members seemed to facilitate and sustain the technology integration change process by involving in various activities. Therefore, a theme “formation of a sustained shared leadership team” was derived from these data analysis processes.

This study adopted triangulation, member check, peer debriefing, and thick description to ensure trustworthiness had been met (Lincoln & Guba, 1985). Multiple data sources, including interview transcripts and document reviews, were compared to confirm the emerging findings. In addition, interview transcripts and document reviews were brought back to the persons who participated to see if the interpretation was correct and the results were plausible. In peer debriefing, the two researchers consistently shared the emerging findings with each other and explained our thinking process to each other to solve disagreement. In other words, the researchers made explicit the decision process regarding how to interpret the data and generated the findings. The researchers provided very detailed descriptive data to support readers’ decision making.

Dependability and confirmability refer to the extent to which the research process can be replicated. The first researcher kept a reflective journal on a daily basis at the stages of data collection and data analysis. The journal consisted of (1) the daily schedule and logistics of the study, (2) a diary that provided opportunities for reflection on

both researchers' values, interests, and biases, and (3) a methodological log with methodological decisions and rationales.

## Findings and Discussion

The purpose of this study was to analyze enabling factors that might have helped to facilitate and sustain the technology integration change process in a multi-section elementary science methods course, SCIED 408, from the perspective of change agents (faculty members and course instructors). Figure 1 describes the three major factors as well as their interaction.

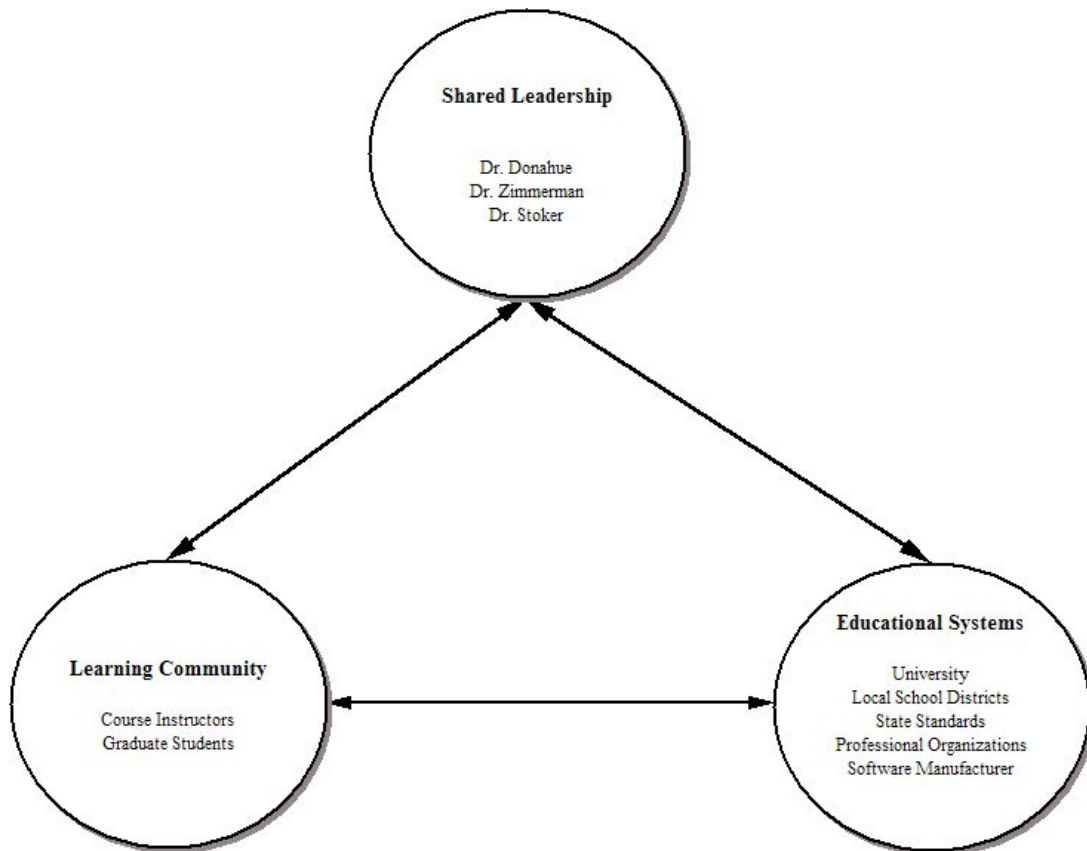


Figure 1. Enabling factors and their interrelationships

### Formation of a sustained shared leadership team

It is essential that leaders possess a number of traits to mobilize and lead change in a technology integration change process. Consistent with the discussion of the important traits of leaders in previous research (Mullen, 2002; Spillane et al., 2004), this study illuminates a number of important traits and elaborates on each of them. These traits include strong expertise, skillful deployment of change strategies, and a shared vision.

First, it is important that leaders possess strong expertise in technology integration to lead change. In this case, Zimmerman's doctorate training placed a heavy emphasis in enhancing preservice teachers' ability in integrating technology in science teaching and learning. Bringing the expertise in technology integration to the science program allowed Zimmerman to enrich the program and continue her interest. Dr. Zimmerman described her enthusiasm in technology integration in science education as she joined the program in 1997:

I came from a program that was very technology intensive. Preservice teachers were expected to teach with technology. I suppose it was only natural to bring that perspective to my work here. ... It

[teaching science with technology] is just very much a part of who I am. So I cannot image that it [teaching science with technology] would not be a fundamental influence on the work that I do.

Donahue had a similar expertise and contributed to technology integration in the program prior to Zimmerman's arrival in 1997. However, Zimmerman's joining stimulated them to pool their efforts and initiate the development of the data collection tools (e.g., Vernier Probes), one of the technology integration projects in the early stage of the change process. Their similar interest and collaboration resulted in the development of their shared vision in technology integration and served as a precursor for large scale technology integration in a later stage of the change process from 1999. Donahue described how he combined his prior efforts with Zimmerman and both of their experiences shaped the development of the technology integration projects. Donahue said:

I had been working on different technology integration projects. Dr. Zimmerman had been working on different technology integration projects. And then she came here in 1997. We both decided that we would combine our efforts.

Ms. Hess said, "I would say the most important persons would be Drs. Donahue and Zimmerman [in early stage of technology integration] ." Ms. Friedman concurred that Donahue and Zimmerman were two major leaders who led the changes in SCIED 408 in the early stage, "I think I would say people [were the critical factors of the changes] . Dr. Donahue and Dr. Zimmerman had higher expectation. And I think that's a big factor." Ms. Aberdeen recognized that Zimmerman and Donahue were major players who initiated the technology integration projects in the early stage.

This finding also reflected that leaders' expertise and learning actions have a heavy impact on the leadership behaviors. In their survey study of 312 respondents consisting of mid-level managers and working professionals across the high-technology corporations and managers enrolled in an executive MBA program, Brown and Posner (2001) found that respondents who reported using more frequently any one of the learning tactics (action, thinking, feeling, and accessing others) also reported engaging more frequently in leadership behaviors like challenging, inspiring, enabling, modeling, and encouraging.

Second, it is also critical that leaders deploy skillful change strategies to identify possible human and financial resources. For example, Donahue, Zimmerman and Friedman co-authored a technology integration state grant proposal in 1998. When the grant was awarded in 1999, the grant supported the hiring of Stoker, the purchase of technology tools, such as PASCO probes, and the revision of the course design. Stoker was hired to be a liaison between the program and the local school districts to facilitate preservice teachers' ability in technology integration in schools. She also built partnerships with these schools and coordinated the STEPS Days activity that invited the elementary students to university classrooms. Zimmerman said:

Typically there was a big disconnect there. Having someone like Dr. Stoker as facilitator allowed us to do what was needed to make teaching with technology happen. For example, we packed up 4 iMacs and took them into schools along with a set of probes. Or we helped students locate the resources in their school districts. Dr. Stoker was the contact person. People would call. She was able to explain and mediate this process of teaching in school. We never had that before. ...There was no excuse for preservice teachers not teaching with technology.

Stoker said: "I put out this blanket: Welcome, please come. We had people from Lewis, Beaver, Boise, Pearson (all pseudonyms) that bring their children, their classes to us." Aberdeen identified the positive influence of the STEPS Days activity that Stoker coordinated. Ms. Hess provided a similar account of the positive influence on preservice teachers' ability to integrate technology and stressed that Stoker was the major person who built this activity.

Skillful change strategies involve more than identifying human and financial resources; encouraging rapport and trust among change agents and stakeholders becomes an important part of leadership. For example, Zimmerman and Donahue brought Stoker along in the change process. Dr. Stoker said:

I always looked at Drs. Donahue and Zimmerman for leadership because they were here beforehand....Drs. Donahue and Zimmerman had the vision, the foresight to realize technology integration in SCIED 408.



After Donahue's role was gradually shifted to the college level and Zimmerman's role became a lead researcher in the context of technology integration, Stoker looked to Zimmerman and Donahue for their prior experience in technology integration from early years. Donahue described the change of his role in SCIED 408 since 1999:

I am not as involved anymore because my jobs have shifted. ...I am able to help other faculty members in other areas to integrate technology, and try to persuade the college what to do so that helps all the students. Not just small numbers of students. Things like that. So I think my role is changed over time.

Stoker built on her experience and then took over the job of being a coordinator of the course, serving as a mentor for new course instructors. Mr. Bell described Stoker's other role as a course coordinator and instructor, "Two years ago, new instructors of SCIED 408 sat in Dr. Stoker's class and watched out how she teaches the course. New people copied off from her and used her ideas." Donahue explicated the stable leadership team formed in the process, particularly the involvement of Stoker:

I would say the number one stimulus was faculty members. ...One is that stable people have been there, which have allowed the notion of integration to continue. Dr. Stoker, as an example, she was there in the beginning and she is still there now. She understands what we are trying to do with that [technology integration] .

Although Stoker's role had changed over time, her stable and consistent involvement in SCIED 408 from 1999 to 2003 enabled her to develop more insights into how to better integrate technology in the course, and she commented on the importance of a stable, long-term perspective:

...Like in your classroom, by the time in your third or fourth year, you are really starting to look more closely at what is helping children learn. ...This is my class, and I am really trying to improve it....This is my most important thing - technology integration in SCIED 408 - to do today, and I am going to do it well. So I guess it is taking more pride, or more ownership, or more concern for what is really happening in the course.

Meanwhile, she learned from Zimmerman's research experience and led SCIED 408 to continuous improvement. Stoker stated:

I try to touch base with Dr. Zimmerman and ask, "What can you tell me from your experiences?" ...So I looked at her for some guidance, especially from her research regarding how I can use that to improve this course.

Such collaboration and trust enabled the successful transfer of knowledge and vision among leadership members. Keeping the three major leaders throughout the project's life enhanced such bonding and trust.

Third, vision is a powerful trait of effective leadership. In this case, shared vision appeared to be a significant factor in sustaining the technology integration change process. A shared vision was developed by Zimmerman and Donahue because of their passion, enthusiasm, and stubbornness in technology integration in early years of the change process. A shared vision was diffused to Zimmerman and Stoker. Although Stoker initially seemed to assume the role of a follower, she built on Zimmerman and Donahue's early experience and became an active and stable participant in improving SCIED 408. A shared vision was evolved to a degree that faculty leaders, course instructors and graduate students viewed technology integration as a visionary direction to follow. As Hess described, they had to do it and there was no way that they could stop doing it. When the people joined the change process in later years of the change process, they viewed technology integration as a positive practice. Therefore, such influence encouraged people to bring more ideas of technology integration into this course, such as Bell's interest in building an online collaborative learning environment to facilitate scientific inquiry.

These findings supported characteristics and traits that are attributed to transformational leaders (Bennis & Nanus, 2003; Mount & Barrick, 1998). In their meta-analysis study of examining the relationship between the Big Five traits and the performance in different types of jobs, Mount and Barrick indicated that Conscientiousness, Extraversion and Openness to Experience correlated positively with job performance because individuals who are dependable, persistent, goal-directed, organized, active, and sociable tend to be high performers. Their findings also demonstrated that Emotional Stability, Agreeableness, and Conscientiousness are indicators of high performance in types of jobs that require interactions with other people because individuals with these traits tend to be more confident, independent, flexible, cooperative, and caring. As evident in this case, Zimmerman, Donahue, and Stoker had a

combination of these traits described above: persistent, goal-directed, confident, independent, cooperative, caring, and encouraging.

A shared leadership team in systemic change is necessary and productive (Lambert, 2002; Reigeluth, 1994, Rowland, 2004). Consistent with the discussions of the role of a leadership team in a change process in a number of studies (Nesbit et al., 2001; Spillane et al., 2004), this case shows that a division among the functions of design, administration, and liaisons among leaders and further illuminates how different expertise and change management strategies needed to lead change were distributed across different leaders. It is evident that a leadership team rather than the individual leader is more appropriate in leading change in this case.

### **Formation of a learning community**

In the technology integration change process, it is a powerful force when these change agents and stakeholders put their expertise together as evident by the learning community in this case. Consistent with the discussions of the importance of a learning community in a change process in previous research (Andrews & Lewis, 2002; Joseph et al., 2002; Nesbit et al., 2001), this study indicates that the formation of a learning community bolstered the change efforts, sustained the change process, and suggested a number of ways to form a learning community.

The first activity was to identify a three-phase conceptual change model to lead the technology integration change process. Zimmerman, along with SCIED 408 course instructors and graduate students in the program, synthesized a three-phase conceptual change model for technology integration that guided the implementation of technology integration in SCIED 408 since 1999. Zimmerman added that the model was valuable in providing the preservice teachers with opportunities to reflect on integrating technology into science learning and teaching in a circular cycle. She said:

Faculty members operated from the conceptual model of thinking about how people learn to teach in general, specifically how they learn to support children's science learning and then what role technology has to play within that. The model where preservice teachers start as science learners and progress to independent teaching over different phases was useful. In guiding our discussions and planning, it is very heavy on reflection in terms of their own practices and their own learning over time.

Donahue added that this model served as a basis for the design of the course and facilitated the implementation of technology integration in SCIED 408:

We design the course in order to have opportunities in that course to support technology integration. And that is the process we found very productive. I should say that learning to be a teacher who uses technology tools is a new concept of teaching. So we needed to support that concept of teaching throughout their course. I think that model was one of the sustaining factors because we all believe in that.

Friedman also stressed that "the technology integration conceptual change model draws the directions. People paid attention to the scaffolding in the model." Stoker pointed out that the technology integration change model is "an excellent way of helping preservice teachers to see how they should use technology to enhance learning. The model continues to be important." Hess also said that "the model drove the organization and served as a way to support preservice teachers in teaching science with technology."

The second activity was collaboration on a technology integration state grant. Zimmerman, Donahue, and Friedman co-authored a technology integration state level grant proposal in 1998 based on the model discussed above, and the grant was awarded in 1999. A number of supportive elements of the grant facilitated the changes, including revision of the course, the ability to buy appropriate technology tools, and the ability to hire key personnel to support important functions. Zimmerman described:

The state grant was instrumental in many ways. ...The grant provided resources that allowed us to make use of cutting edge technologies in a systematic way.

Hess added:

The grant was really the stimulus for implementing the technology. We were able to buy a lot of the technology and modify the course to be technology intensive....When I first started in 1998, we did

not have STEPS Days. STEPS Days was the result of the technology integration grant. So part of it we wanted them to have opportunities to teach with technology and guided environment which would be on campus before the preservice teachers went out to the classroom and try to do it with the students. That is the idea.

Stoker also said,

When I came here on the technology integration grant in 1999, I worked with many elementary schools. Whenever preservice teachers would be going for their field experience and wanted to use the probes, Science Court, or did not have computer in the school, I was holding computers or software out there for them to be able to use in the field experience classroom.

The third activity was the establishment of a research team. Zimmerman and Donahue conducted research in the context of this course with the instructors and graduate students, including Friedman, Hess, Aberdeen, and Bell. Friedman called it “the research community.” Working with Zimmerman, Friedman, Hess, and Aberdeen’s research interest centered on the influence of various science specific tools on science inquiry and reflections. Bell’s research focused on the perceptions of elementary preservice teachers of science inquiry in an online collaborative learning environment. Zimmerman emphasized the role of research findings in providing evidence to support the proposition that technology was a powerful tool in enhancing science learning and teaching:

It is the importance of our own research, our own practice in the context of 408 in terms of how we are modifying and changing assignments over time because we are constantly learning from that process.... Various components are sustained only to the extent that they fit with the model. If we do not have evidence supporting how our preservice teachers are learning to teach science, then we would not keep it around. But what tends to happen is we get evidence of how we might help do things better. So certain projects get modified or changed. But the technology piece has not gone away because it is pretty powerful.

The fourth activity focused on the improvement of teaching. The course instructors held regular meetings to discuss improvements. A few veteran instructors served as mentors for new instructors. Stoker described how she worked with other instructors on a regular basis and how she modeled teaching practices for new instructors:

We hold weekly meetings. I put my lesson plans up in my web site for new instructors to pull down. Then we revise those together and come up with ways to improve. That’s what I brought about little problem-based learning groups. I guess I lay out the ideas that we need to cover and I assign the readings and so on. But then we come up with how to teach the class and work.

Hess emphasized why the change process had sustained:

I think the big part of that is the team approach. It is a team that teachers work very closely together. So, it is not like each individual does their own thing. We follow the same syllabus....

Zimmerman added, “There has to be a bigger commitment of the team in the process.” Bell also illustrated his experience of developing an online collaborative learning environment for other course instructors in recent years:

I have become expert on this online collaborative learning environment because I spent most time with it. So in the fall semester of 2002, I was on everyone’s site as an administrator. I was the one to whom other instructors came for questions. So I think my role has increased their confidence.

### **Influences from different educational systems**

In investigating the technology integration change process, it is critical to look beyond the course to other educational systems with which the course interacts. It is also important to identify the shifts and changes at different levels of educational systems that may influence technology integration at another level. Previous research identified the influences from local school districts (Taylor & Wochenske, 2001), the university level (Grove, Strudler, & Odell, 2004), and the state level and software vendors (Brooks-Young, 2007; Brown, 2007); this study shows similar findings and illuminates more dynamics within each of the educational systems and explicates the interactions among them. First, although the university’s support for technology integration was not solely intended for SCIED 408, the changes came at a time when it was most supportive of change in SCIED 408. In terms of online collaborative tools, the university offered a university-wide course management tool, A New Global Environment for Learning (ANGEL), as a common instructional tool for the instructors and the students. Before, the course used tools

such as Creating Communities of Learners (CCL) that were offered outside the university, and which failed to continue because of complicated logistical and technical problems. Donahue explained:

Mainly CCL was not housed in our server. So the maintenance issues associated with that were becoming more complicated. We wanted the permissions to add or change students. We had to go through too many layers of people to do that. So the main reason is because we did not have as much control over the environment.

At the school district level, all participants repeatedly identified partnerships between SCIED 408 and the local school districts through the STEPS Days activity as a positive factor of sustaining the technology integration change process. Donahue also stressed the importance of establishing the partnerships with local school districts by Stoker:

Dr. Stoker's job turned out to negotiate with all the teachers and technology coordinators and others in the school districts. So students can integrate technology into their teaching. We had a loan at that time. ...So our students more or less were able to integrate technology into lessons without any difficulty.

The science academic standards from the state department of education also played a role in dictating the design of the course and drove the direction of the types of technology integrated into the curriculum. Bell said, "We wanted to do something that was environmental to reflect the environmental ecology standard of the State. Probably the sound probe works better for what we want to accomplish." Stoker and Aberdeen added, "The new assignment is on the watersheds. And that lends itself nicely to using the probes of temperature and pH. It is one of the new standards from environmental ecology."

Participants regarded that political attention from professional organizations was a big drive for the technology integration change process. Zimmerman, Donahue, Aberdeen, Hess, Friedman, and Bell had been very productive in publishing journal articles and presenting conference papers on technology integration into the SCIED 408 course, which drew the attention and interest of people working in similar areas of research and their accomplishment was reinforced. The program also received recognition from the National Council for Accreditation of Teacher Education (NCATE). Hess claimed that such attention sustained their efforts in the change process. She said:

When NCATE came in and did accreditation reviews, one certain thing [technology integration] that they hailed....The science education program was getting attention, positive attention. They were related to the technology. So it almost became something that we cannot quit now. So we need to keep striving. We need to keep trying to do better. There was political attention that, I think, helped sustain it.

The partnerships between SCIED 408 and the software manufacturer sustained technology integration. The software development manufacturer, PASCO, provided more tools and professional training for the faculty members and the instructors. Donahue explained:

We bought some. And PASCO gave us more. And over the years, people like Dr. Stoker wound up getting involved with PASCO and became PASCO technology educator. Then they go up and do workshops, which I think is great.

The partnerships provided numerous opportunities for the faculty and course instructors to grow and develop new expertise. Stoker described:

One way that I try to keep up with is I became PASCO technology educator, which I went out for training. When I go to conferences, I work with those folks at their booth. Selling probes. (Laughter) I am not really selling but bringing folks over and demonstrating the use of software. ...In fact, we are trying to arrange for one of the sales reps to come to us and do a demonstration to bring the new graduate students onboard with the technology.

The manufacturer also accommodated the technology tools to elementary level in recent years. Zimmerman added, "So the PASCO developed a USB probe, which is much more elementary friendly. These tools for sciencing get more accessible for the elementary classroom."

It is noted that the links among the leadership team, the learning community, and educational systems facilitate the development of change agents' and stakeholders' expertise as well as help expand the resources available to the course. In this case, Stoker became the PASCO technology educator and worked with the manufacturer to share her experience with other schools at conferences. The software manufacturer also provided professional training for

course instructors and graduate students and improved their software geared more toward the elementary level. The benefit is reciprocal rather than one dimensional.

We have identified three enabling factors and described their interaction. These three factors complemented and enhanced each other. The three factors found in this study were similar to a number of transformational variables, such as external environment, leadership and individual and organizational performance, described in Burke and Litwin's (1992) organizational performance and change model. They identified that these variables have causal relationship among them and argued whether the change is caused by the forces of the external environment or leadership. In this case, it seems that the technology integration change was pushed by the university faculty downward. Based on the findings, we argued that the university faculty is in the best position to champion a technological change initiative and is in control of leading the technology integration change process.

## **Implications and Conclusions**

The ultimate goal of this study was to provide a framework of technology integration for persons who play the role of change agents in educational settings. This section summarizes the findings and addresses the implications for this study. Additionally, I provide recommendations for further study.

As indicated in this study, continuing commitment of a transformational leadership team sustains the change. The transformational leadership style in this study truly reflects Senge's and Lannon-Kim's (1991) view on current view of leadership in systemic change. Senge and Lannon-Kim (1991) stated, "In systems approach, leaders' roles will differ dramatically from that of the charismatic decision-maker. They will be designers, teachers, and stewards." Senge and Lannon-Kim (1991) further explained that "leaders' sense of stewardship operates on two levels: stewardship for the people they lead and stewardship for the larger purpose or mission that underlines the enterprise." Thus educational reformers should adopt a similar leadership style. That is, it is essential to look for the persons at higher levels within the educational institution who hold the vision of technology integration, have commitment to a variety of change activities such as design, and are open to share ideas and beliefs with other capable subordinates. Then form a group at the outset of the change process and continue to work with each other to deal with issues that emerge throughout the change process.

The formation of a learning community serves as powerful strategy to facilitate and sustain technology integration. Research conducted by Balach and Szymanski (2003) investigated the dynamics of a learning community through collaborative action research. This study reveals that the dynamics of collaboration, inquiry, parity, reflective dialogue and shared vision shape the progression of the group. More specifically, the research community in this case truly reflects these characteristics. A research community is one type of a learning community because they strive to improve students' performance by involving them in research activities.

It is necessary to identify different levels within the educational system that place influence on the change process. This allows for adopting approaches to strengthen the synergy among these systems and tackle the problems that may arise from these systems. For instance, at the local community level, it is crucial to involve local school districts' teachers, students, and parents. It is essential to plan a variety of ways to establish relationships with the community, such as delegating a particular person to coordinate activities that involve local community and local school and district people on a regular basis. Additionally, the focus could be on developing partnerships with software manufacturers for appropriate support. The companies can provide technical assistance and buy-in. They can also offer professional development workshops for the potential users and learners.

There are a number of implications for persons who play the role of change agents. First, it is appropriate to provide training in leadership skills and change strategies for persons who lead the change. Specifically, the skills include identifying other 2 or 3 persons who might share the same vision and then assigning different responsibilities based on areas of expertise. Second, it is critical to inform persons who lead change of different leadership styles and model the effective leadership behaviors. Third, in order to bring collaboration among change agents and stakeholders, it is important to design activities at different stages of a technology integration change process and allocate specific, regular time for these activities. Fourth, it might be beneficial to assist persons who lead change and stakeholders in developing strategies in identifying the potential positive and negative influences from other educational systems and examine ways to combat or enhance these influences.

In addition, I would like to provide recommendations for future study. The first recommendation is to conduct a longitudinal study in sustenance of change. Holloway (1996) critiqued that most studies on the technology integration change lacked longitudinal study. This study was intended to overcome these criticisms and focused on a six-year case to illuminate the factors and issues involved in the technology integration change process in educational setting. Thus, a lengthy period of study on examining factors that affect the change process is recommended.

While the findings are limited to the perspectives of change agents, the second recommendation is to research on the perspectives from other stakeholders such as preservice teachers and school teachers in order to gain a holistic, systemic picture of a technology integration change process. The holistic, systemic view of a change process enables the people involved in change efforts to plan for the establishment of roles, structures, and resources to bring the people and educational systems together to facilitate and sustain technology integration in a long term change effort.

## References

- Andrews, D., & Lewis, M. (2002). The experience of a professional community: teachers developing a new image of themselves and their workplace. *Educational Research*, 44(3), 237-254.
- Balach, C. A., & Szymanski, G. J. (2003). *The growth of a professional learning community through collaborative action research*. A paper presented at the 2003 annual meeting of the American Educational Research Association, Chicago, IL.
- Banathy, B. H. (1991). *Systems design of education: A journey to create the future*, Englewood Cliffs, NJ: Educational Technology.
- Banathy, B. H. (2001). We enter the twenty-first century with schooling designed in the nineteenth. *Systems Research and Behavioral Science*, 18, 287-290.
- Bennis, W. G., & Nanus, B. (2003). *Leaders: strategies for taking charge*. New York, NY: Harper Business.
- Brown, K. (2007). *Technology Funding: A How-To Guide*, retrieved August 2, 2008 from <http://www.techlearning.com/story/showArticle.php?articleID=196604546>.
- Brown, L. M., & Posner, B. Z. (2001). Exploring the relationship between learning and leadership. *Leadership and Organization Development Journal*, 22 (5/6), 274-281.
- Brooks-Young, S. (2007). Going Corporate: Corporations Can Be a Great Source for Grant Funding. *Technology and Learning*, 27(7), 32-36.
- Brush, T., Glazewski, K., & Rutowski, K. (2003). Integrating technology in a field-based teacher training program: The PT3@ASU project. *Educational Technology Research and Development*, 51(1), 57-72.
- Burke, W. W., & Litwin, G. H. (1992). A causal model of organizational performance and change. *Journal of Management*, 18(3), 525-545.
- Cibulka, J., & Nakayama, M. (2001). *Practitioners guide to learning communities*, Washington, DC: National partnerships for Excellence and Accountability in Teaching.
- Dexter, R. R. (2005). Instructional leadership for systemic change: The story of San Diego's reform. *School Administrator*, 62(9), 57-57.
- Fullan, M. (2007). *The new meaning of educational change* (4<sup>th</sup> Ed.), New York, NY: Columbia University.
- Havelock, R., & Zlotolow, S. (1995). *The change agent's guide* (2<sup>nd</sup> Ed.), Englewood Cliffs, NJ: Educational Technology.
- Holloway, R. E. (1996). Diffusion and adoption of educational technology: A critique of research design. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology*, New York, NY: Simon & Schuster Macmillan, 1107-1133.
- Grove, K., Strudler, N., & Odell, S. (2004). Mentoring toward technology use: Cooperating teacher practice in supporting student teachers. *Journal of Research on Technology in Education*, 37(1), 85-109.
- Joseph, R., Jenlink, P. M., Reigeluth, C. M., Carr-Chellman, A., & Nelson, L. M. (2002). Banathy's influence on the guidance system for transforming education. *World Future*, 58, 379-394.

- Lambert, L. (2002). A framework for shared leadership. *Educational leadership*, 59(8), 37-40.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*, Beverly Hills, CA: Sage.
- Linn, M. C. (2003). Technology and science education: starting points, research programs, and trends. *International Journal of Science Education*, 25(6), 727-758.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- Miles, M., & Huberman, M. (1994). *An expanded sourcebook: Qualitative data analysis*. Beverly Hills, CA: Sage.
- Mount, M. K., & Barrick, M. R. (1998). Five reasons why the "Big Five" article has been frequently cited. *Personnel Psychology*, 51(4), 849-858.
- Mullen, C. A. (2002). The original ten: a multisite case study of Florida's Millennium High School reform model. *Education policy analysis archives*, 10 (40), Retrieved May 20, 2008 from <http://epaa.asu.edu/epaa/v10n40.html>.
- National Research Council (NRC). (2001). *Educating teachers of science, mathematics, and technology: New practices for the new millennium*, Washington, DC: National Academy Press.
- Nesbit, C. R., DiBiase, W. J., Miller, A. S., & Wallace, J. D. (2001). *In their own words: what science and mathematics teacher leaders say are important aspects of professional development*, Chapel Hill, NC: University of North Carolina at Chapel Hill.
- PASCO Scientific (2008). *PASCO Probeware (Version 1) [Computer software]*, Roseville, CL: PASCO Scientific.
- Reigeluth, C. M. (1994). The imperative for systemic change. Envisioning a new system of education. In C. Reigeluth, & R. Garfinkle (Eds.), *Systemic change in education*, Englewood Cliffs, NJ: Educational Technology, 3-11.
- Rowland, G. (2004). The concept of sustainability in the evolutionary guidance of an educational institution. *Systemic Practice and Action Research*, 17(4), 285-296.
- Seels, B., Campbell, S., & Talsma, V. (2003). Supporting excellence in technology through communities of learners. *Educational Technology Research and Development*, 51(1), 91-104.
- Senge, P., & Lannon-Kim C. (1991). Recapturing the spirit of learning through a systems approach. *The School Administrator*, 3, 23-28.
- Spillane, J. P., Diamond, J. B., Walker, L. J., Halverson, R., & Jita, L. (2001). Urban school leadership for elementary science instruction: identifying and activating resources in an undervalued school subject. *Journal of Research in Science Teaching*, 38(8), 918-940.
- Stake, R. E. (1995). *The art of case study research*, Thousand Oaks, CA: Sage.
- Surry, D. W., & Ely, D. P. (2006). Adoption, diffusion, implementation, and institutionalization of instructional design and technology. In R. A. Reiser, & V. J. Dempsey (Eds.), *Trends and issues in instructional design and technology*, New Jersey, NJ: Pearson Education, 104-112.
- Taylor, N., & Wochenske, J. (2001). The San Diego Science Alliance. In Thorson, A. (Ed.), *Partnerships with Business and the Community*, Columbus, OH: Eisenhower National Clearing House for Mathematics and Science Education, 37-38.
- Tom Snyder Productions (2008). *Tom Snyder Science Court (Version 1) [Computer software]*, Watertown, MA: Tom Snyder Productions.
- Vernier Software Technology (2008). *PASPORT Sensor (Version 1) [Computer software]*, Beaverton, OR: Vernier Software Technology.
- Windschitl, M. (2000). Supporting the development of science inquiry skills with special classes of software. *Educational Technology, Research, and Development*, 48 (2), 81-96.
- Yin, R. (2002). *Case study research: Design and methods* (3<sup>rd</sup> Ed.), Newbury Park, CA: Sage.
- Zemba-Saul, C., Munford, D., & Friedrichsen, P. (2002). Technology tools for supporting scientific inquiry: A preservice science education course. *Proceedings of the 2002 Annual International Conference of the Association for the Education of Teachers of Science*, ERIC Document Reproduction Service No. 465 602.

## Appendix A: Interview Protocol

### Role of Participant

1. How long have you been in SCIED 408 as a course instructor?
2. How has your role changed over time in terms of different technology integration projects in SCIED 408?
3. How did you gain support from others who played different roles (e.g., other course instructors) in the change process?

### Course

4. Could you give me a description of SCIED 408?

### Technology Integration

5. Who initiated the ideas of integrating science specific technology tools into SCIED 408?
6. Could you describe current technology integration projects in SCIED 408, such as major technology activities and assignments?
7. Could you describe the history of each project (e.g., when was it started; why was it initiated or motivated; how has it changed)?
8. Are there any technology integration projects in SCIED 408 that were dropped and are not listed on current list? Why?
9. How have elementary preservice teachers' perceptions about their roles and about technology integration in elementary science education changed (e.g., evidence)?

(Thank individual for participating in the interview. Assure him or her of confidentiality of responses and potential future interviews.)

## Appendix B: Time-ordered Matrix of Technology Initiatives

	Time Periods						
	Before 97-98	97-98 Year 1	98-99 Year 2	99-00 Year 3	00-01 Year 4	01-02 Year 5	02-03 Year 6
Technology Integration Projects	Data collection tools (Vernier probes)	Data collection tools (Vernier probes)	Data collection tools (Vernier probes)	Data collection tools (PASCO probeware)			Data collection tools (PASCO probeware)
		Online collaborative tools (Connecting Communities of Learners)	Online collaborative tools (CCL)	Online collaborative tools (ANGEL)			Online collaborative tools (ANGEL)
				Simulations and modeling tools (Tom Snyder Science Court)			Simulations and modeling tools (Tom Snyder Science Court)



### Appendix C: The Roles of Participants

	Before 97-98	97-98 Year 1	98-99 Year 2	99-00 Year 3	00-01 Year 4	01-02 Year 5	02-03 Year 6
Zimmerman		Course instructor  Course coordinator  Lead researcher	Course instructor  Course coordinator  Lead researcher	Course coordinator  Lead researcher		Lead researcher	
Donahue	Course instructor	Course instructor	Coordinator of Science Teacher Education Program				
Stoker				Course instructor  Liaison	Course instructor  Liaison	Course instructor  Course coordinator  Liaison	
Hess		Course instructor				Graduated	
Friedman			Co-authored the state level grant proposal with Donahue and Zimmerman	Research associate			Graduated
Aberdeen						Course instructor	
Bell						Course instructor	

## Teachers' Perceptions of the Dimensions and Implementation of Technology Leadership of Principals in Taiwanese Elementary Schools

I-Hua Chang, Joseph M. Chin

Department of Education, National Chengchi University, No. 64, Sec. 2, ZhiNan Rd., Taipei City 11605, Taiwan // Tel: +886-2-29387738 // Fax: +886-2-29387737 // ihchang@nccu.edu.tw // mcchin@nccu.edu.tw

Cheng-Mei Hsu

Graduate School of Design Doctoral Program, National Yunlin University of Science and Technology, No. 123, Sec. 3, University Rd., Douliou, Yunlin 64002, Taiwan // Tel: +886-2- 86613915 // Fax: +886-2-29609046 // g9330811@yuntech.edu.tw

### ABSTRACT

Principals' technology leadership is strongly correlated with teachers' integration of educational technology, and technology leadership is necessary for effective utilization of technology in schooling. The article describes a study that empirically investigated teachers' perceptions of elementary schools principals' technology leadership practice in seven cities in Taiwan for understanding the implementation of technology leadership. This study used structural equation modeling with prospective data to test for model fit. The findings identified the four constructs (i.e., vision, staff development, infrastructure support, evaluation and research) comprising principals' technology leadership. The findings also show that interpersonal and communication skills are important antecedents to principals' overall effective technology leadership. Four themes (e.g., budget shortage, technology facilities, staff development, and leadership problems) which emerged from the transcript data were the practical problems that principals faced while they implemented technology leadership in their schools. The results suggest that principals who embrace technology will effectively lead their schools to acquire educational resources to enhance student engagement and learning.

### Keywords

Elementary schools, Principal technology leadership, Technology integration, Technology literacy, Technology vision

### Introduction

Technology development has historically facilitated progressive human civilization, improved living environments, and increased human welfare (Shen, 2004). With information technology development and innovation, computers, the Internet, and other information technologies are becoming important learning tools in students' everyday lives. Campus information technology utilization is designed to help students and improve educational quality. Therefore, developing student technology literacy is becoming increasingly important. Principals should possess basic information technology skills and literacy (Scott, 2005; Wexler, 1996) to support staff and faculty in preparing students to face information-age challenges. Technology Leadership Academies have been established in every U.S. state administrative office to provide curriculum projects for principals and administrators to stay in step with flourishing information technology development. In the modern information explosion environment, technology education becomes increasingly vital day by day, and principals with efficient technology leadership skills are the key to successful policies and technology education plans (Chang & Tseng, 2005).

The emerging technology leadership role means that principals cannot ignore campus technology management. Assuming a technology leadership role entails promoting technology literacy to prepare students for the information age. Principals' new leadership roles are becoming increasingly important in schools. Ross and Bailey (1996) indicate that as leaders who lay the educational foundation for their schools, principals have quickly become leaders who promote and support new educational technologies. More than ever, they are acting as facilitators of change who pursue new technological advancements and innovations that may benefit student achievement and learning. Thus, the principal's role becomes crucial in efforts to acquire and implement new educational technologies within [public] school settings.

Reeves (2004) conducted the National Leadership Evaluation Study from March 2002 to September 2002 with a nonrandom sample of 510 leaders, including district superintendents, central office administrators, and principals

from twenty-one U.S. states. The major dimensions for constructive leadership evaluation in his study included technology, faculty development, leadership development, and learning. The technology dimension consists of demonstrating the use of technology to improve teaching and learning, personal proficiency in electronic communication, and coherent management of technology resources, technology staff, and information. Based on the dimension and content of the technology mentioned, it is evident that school principals and administrators should pay attention to the technology issue.

The role of the principal has shifted from a narrow focus on management to a broader scope of leading student learning, reflecting the vision of building, facilitating, and supporting practices of leadership to create change and continual educational improvement in accountability-defined arenas (as cited in Orr & Barber, 2006). The dramatic change of the principal's role since the early 1980s has evolved from being primarily a building manager (Sharp & Walter, 1994), to an instructional and curriculum leader (Checkley, 2000; Cheng, 2004; Glatthorn, 2000; Huang, 2004; Wu, 2004), and more recently to a technology leader (Anderson & Dexter, 2005; Bailey & Lumley, 1994; Chang, 2002; Chang, 2003a, 2003b; Ford, 2000; Hsieh, 2004; Inkster, 1998; Kadela, 2002; Matthews, 2002; Ross & Bailey, 1996; Scott, 2005; Seay, 2004; Stegall, 1998; Yeh, 2003). New technology-related standards and performance indicators (e.g., leadership and vision; learning and teaching; productivity and professional practice; support, management, and operations; assessment and evaluation) for administrators have been developed, and principals' technology leadership roles have been explored as a means of improving student performance and supporting effective integration of technology into schools (Bridges, 2003; Hughes & Zachariah, 2001; ISTE, 2005; Kadela, 2002; Matthews, 2002; Seay, 2004).

Prior to relevant technology leadership research, which is gradually emerging (Battle, 2004; Bridges, 2003; Calhoun, 2004; Frazier, 2003; Hudanich, 2002; Martens, 2003; Mirra, 2004; Nash, 2002; Rogers, 2000; Scanga, 2004; Scott, 2005; Shuldsman, 2003; Ury, 2003; Wagner, 2004; White, 2004), leadership theory evolved over the decades from trait theory, behavior styles theory, situational theory, and transformational theory to a new leadership paradigm. To cope with the leadership paradigm shift, scholars proposed new school leadership strategies including technology leadership, structure leadership, moral leadership, cultural leadership, symbolic leadership, human resource leadership, political leadership, and strategic leadership. The main responsibility of technology leadership, for example, is to identify the connections among technology, school vision, school mission, and educational policy. In other words, school leaders should understand the importance of computer and information technology for students as well as enrich the technology environment for student learning. Meanwhile, school leaders should empower, encourage, and collaborate with experts and local businesses to support campus-wide technology infrastructure (Chang, 2005). Information age school leaders should model technology leadership behaviors that promote teaching and learning to foster a learning environment in their organization. Principals should play a critical technology leadership role to create a learning culture.

Published in April 1983, *A Nation at Risk* called for reforms throughout the nation's K-12 educational system. The report specifically emphasized the need for students to become more technologically literate. Since then, a proliferation of educational technology policies and implementation plans has been launched in K-12 schools (Aten, 1996; U.S. Department of Education, 2001). Three prominent national reports serve as examples to illustrate these policy changes: (1) *The Goals 2000: Educate America Act*, placing educational technology in a prominent position (Glennan & Melmed, 1996); (2) *Getting America's Students Ready for the 21st Century: Meeting the Technology Literacy Challenge*, releasing the nation's first educational technology plan (ED Office of Educational Technology, 2001); and (3) *Enhancing Education Through Technology in No Child Left Behind* (2001), unveiling a new educational technology plan (No Child Left Behind, 2001). These reports emphasized the importance of and need for educational technology in public education and suggested that with new technology policy implementation, the next generation of American children would be well prepared for the information age. Similarly, in Taiwan, information technology policy has been integrated into instruction in schools via programs that have been in place for a decade. In light of the importance of technology integration into school curricula and instruction, numerous technology literacy training initiatives have been implemented for teachers in Taiwan during the past decade. Additionally, technology literacy training programs and workshops have been offered to school administrators such as elementary and secondary school principals in Taipei City, Taichung City, and Kaohsiung City. Therefore, it is critical to investigate and evaluate the effectiveness of technology integration due to the government's huge investment in the programs mentioned above.

British research regarding technology leadership may be represented by Robinson (1994). According to Robinson, school administrators should support teachers in understanding the potential of technology while identifying applicable software and hardware. To this end, administrators ought to facilitate the exchange of ideas regarding uses of information technology through team teaching, the creation of work teams, work development checklists, and other resources or methods. Administrators should also evaluate the effectiveness of staff members in using information technology in order to identify staff members who require additional professional training. To provide practical training to principals of middle and elementary schools, the Hong Kong Education and Manpower Bureau designed a course to nurture future information technology leaders. The goal of the course was to foster the development of principals' knowledge and skills, while also helping them to understand the influence of information technology on pedagogy. The curriculum encouraged principals to initiate "a vision for information technology education in the school," understand "information technology leadership and organizing for strategic information technology," appraise and optimize "priorities in developing the school's information technology strategies," develop "a culture of collaborative learning through the internet" and "an edifying learning community, supporting the principal in integrating and utilizing appropriate information and communication technologies in pioneering novel pedagogies"(para. 1). At the same time, the course used academic focus groups to equip middle and elementary school principals with the following educational tools after completion of their learning courses: (1) the ability to pioneer a network of information technology education; (2) a profound understanding of the multidimensional aspects of state-of-the-art information technologies; (3) the ability to delineate information technology action plans for the new academic year; (4) a penchant for exploring applicable information technologies and learning/teaching models; (5) through personal observations and debates with others, the ability to utilize learning experiences from the training course to achieve the ultimate goal of reforming the curriculum by integrating information technology to the classroom; (6) an understanding of novel information technologies in the educational context (Centre for Information Technology in Education, 2005).

Technology leadership has been explored in the US since 1990. This body of research indicates that technology leadership has been defined in terms of concepts, indicators, nature, and behavioral characteristics, which emphasize that leadership in technology could be implemented as a function of innovations in schools. Compared to the volume of research regarding technology leadership in the U.S., there is little information detailing technology leadership in Taiwan. With relatively few studies specifically addressing evaluation of principals' technology leadership, this area necessitates future exploration so that current and future leaders can be prepared to deal more effectively with technology and to successfully implement technology policy.

### **The Purpose of the Study**

Studies show that technology leadership can have a significant impact (Anderson & Dexter, 2005; Bridges, 2003; Inkster, 1998; Kozloski, 2006; Rogers, 2000; Weber, 2006; Yoho, 2006). Therefore, the purpose of this study is to empirically investigate teachers' perceptions of principals' technology leadership practices in seven different cities in Taiwan. Three primary questions will be addressed: (1) what are the dimensions of principals' technology leadership, (2) are the dimensions of principals' technology leadership perceived to be important to teachers; that is, do perceptions vary according to teacher demographic characteristics, and (3) what practical problems are faced by principals implementing technology leadership in their schools?

### **Theoretical Framework**

Leadership is a key element for successful educational reform or innovation. For effective technology use in school districts' instructional programs, strong leadership must be provided at both the school and district level (Cory, 1990). Integrating educational technology into classroom instruction requires basic changes to current school models (Kinnaman, 1994), and principals must effectively model technology leadership (Ross, 1993). Principals are the key players in the educational change process (Ross & Bailey, 1996) and need to know the importance of effective school management and improving classroom instruction (MacNeil & Delafield, 1998). Thus, the principal's technology leadership proficiency is paramount to the current needs of public education.

Researchers have suggested that if school leaders are to help their institutions apply technology in beneficial ways, their leadership should: (1) empower the principal's team members (e.g., teachers, staff members), (2) identify the

principal's role during technology integration, (3) understand the interconnectedness and complexity of the principal's technology role, and (4) establish baseline information at the beginning of the principal's technology integration process (e.g., Bailey, 1997; Ford, 2000; Inkster, 1998; Kearsley & Lynch, 1994). Principals who effectively lead technology integration within their schools typically perform well in leadership and management, vision and goal setting, student learning, teaching, professional development and training, operations and infrastructure support, and assessment and evaluation (ISTE, 2001).

Effective leadership is a key element for the success of any educational innovation or new school instructional program (Cory, 1990). Effective principals should be actively involved in all aspects of educational technology (Inkster, 1998). Furthermore, Stegall (1998) suggests that principals' technology leadership is essential in elementary schools. A critical technology leadership element is the ability to develop and articulate a vision of how technology can produce educational change (Kearsley & Lynch, 1994). More importantly, technology leadership skills are necessary for principals to pursue new and emerging educational technologies for their schools (Bailey, 1997).

Recent educational literature is replete with studies related to technology leadership (Anderson & Dexter, 2000; Appalachia Educational Lab, 2000; Bailey, 1997; Brush, 1998; Ferris & Roberts, 1994; Jewell, 1998; Keating, Stanford, Self, & Monniot, 1999; Kowch & Walker, 1996; Robinson, 1994; Thomas & Knezek, 1991). For example, Aten (1996) stated that technology leadership supports effective instructional practices through a combination of interpersonal skills, knowledge of a variety of current technology applications, and the vision to anticipate future technology-based solutions for education. Murphy and Gunter (1997) also suggested that leadership should model and support computer technology to result in more effective curriculum integration of technology by teachers. Kearsley and Lynch (1994) noted that the manner in which technology is implemented is more important than any intrinsic educational process characteristics. They go on to say that technology is a powerful tool that supports school reform and facilitates student learning. The potential benefits of good leadership can include improved academic achievement by students, improved student attendance and reduced attrition, better vocational preparation of students, more efficient administrative operations, and reduced teacher/staff burnout and turnover. More than ever, principals play a critical role in implementing and improving technology education in their schools.

Drawing from the empirical literature on principals' leadership in general and, specifically, their effectiveness as technology leaders, five primary dimensions of principals' technology leadership will be examined and serve as the conceptual framework for this study: vision, planning, and management; staff development and training; technology and infrastructure support; evaluation and research; and interpersonal and communication skills. The aforementioned five dimensions were chosen because they are the principals' core tasks in dealing with teaching and learning as well as administrative operations with technology in their schools.

Vision and planning have been demonstrated to be very important technology leadership characteristics. Effective technology leadership develops and articulates a vision of how technology can produce school change (Cory, 1990). Developing such a vision requires principals to clearly understand district, state, and national trends and movements taking place with new and developing technologies. Inkster (1998) noted that creating a vision of how technology should be used by teachers and students is a significant indicator of a principal's technology leadership. Principals must have a clear technology vision and understand technology implications for the classroom. Without vision, staff members who lack direction and guidance for technology integration will not succeed (Ross & Bailey, 1996). Stakeholders (e.g., parents, community members, teachers, students) must also be involved in the school's technology vision. The greater the stakeholders' consensus and commitment, the more likely it is that the principal's technology vision and planning will become a reality (Jewell, 1998).

Staff development and training are important aspects of technology leadership. The most important responsibility identified by technology leaders was the ability to describe and identify resources for staff development (Ford, 2000). Effective staff training must consist of describing and identifying resources, and planning and customizing development programs based on individual and school needs. For example, the in-service plan should include listings and schedules of technology workshops and courses available to all administrators, educators, and support staff. Curriculum guidelines and effective technology leadership are also critical to the planning and designing of educational staff development activities (ISTE, 1998). To achieve an optimum staff development plan, principals need to identify key resources and players who can provide formal and informal leadership and technology support at every grade level and within every discipline to accomplish an effective instructional technology plan (Moursund, 1992).

Acquiring technology and supporting the infrastructure are crucial areas of technology leadership. Technology leaders need to provide service and technical support to their schools (Bailey, 1997). Principals, as technology leaders, must provide access to and the opportunity to acquire technology resources, as well as ensuring that appropriate facilities for technology are well supported (Collis, 1988). Assisting staff with a variety of issues such as purchasing appropriate software applications, troubleshooting equipment problems, installing equipment and infrastructure, maintaining and repairing equipment, understanding a variety of operating systems, and managing and allocating resources fairly and effectively are desired skills suggested for technology leaders (Aten, 1996; Ford, 2000). Providing and ensuring access to technology and maintaining infrastructure support were two areas most often identified as critical elements of principals' technology-related behaviors (Inkster, 1998).

Evaluation and research should be of primary concern to technology effectiveness. Effective principals implement evaluation procedures that allow for growth assessment of teachers and staff members toward established technology standards and help guide their professional development plans (ISTE, 2001). Principals should also include the learning and teaching process as a criterion in assessing instructional staff performance in the use and application of educational technology (ISTE, 2001). Cory (1990) suggests that because of the rapidly evolving nature of instructional and learning programs, it is particularly important that these programs are evaluated annually and the results incorporated into ongoing and future planning and assessment processes. Effective technology leadership should include evaluations of new and existing technology in terms of cost, benefits, and educational impact (Aten, 1996). Such evaluations provide principals with the appropriate information to effectively assess and improve technology plans in their schools.

Interpersonal and communication skills can impact principals' effective technology leadership. The ability to interact and communicate well is an important technology leadership characteristic (Aten, 1996; Inkster, 1998; Kline, 1993). Leaders must be able to get along with teachers and staff members as they begin to integrate new learning technologies (Bailey & Lumley, 1994; Jewell, 1998). A principal can be an effective leader without technological expertise; however, without interpersonal and communication skills, principals cannot be effective technology leaders. Technology leadership requires refined interpersonal and communication abilities, as well as technological competency (Ray, 1992). Principals' communication skills are often closely tied with their effective technology leadership (Inkster, 1998).

## **Research Method of the Primary Study**

In order to investigate principals' role in facilitating technology use in U.S. schools, the author began to explore technology leadership dimensions in his 2002 doctoral dissertation (Chang, 2002), providing seven dimensions of technology leadership (i.e., vision, planning and management; in-service training; interpersonal and communication skills; ethical and legal issues; integrating technology into curriculum and learning; technological support and infrastructure; evaluation, research and assessment). In 2003, the author modified the previous research to include five dimensions of technology leadership. In his technology leadership model, interpersonal and communication skills are important antecedents to principals' overall effective technology leadership (Chang, 2003b). To further explore technology leadership development in Taiwan, the author revised the *Technology Leadership Questionnaire* consisting of thirty-nine items created in 2003 and applied the revised *Principals' Technology Leadership Questionnaire* to measure 434 teachers' perceptions regarding dimensions and implementation of principals' technology leadership in Taichung City, Taiwan (Chang, 2004). The author revised the *Principals' Technology Leadership Questionnaire* in 2005, expanded the sample size, and randomly sampled 1024 teachers selected from 188 elementary schools in seven cities in Taiwan. This was done to develop appropriate technology dimensions and technology standards for leaders in terms of school management and development. Providing an empirical measure of teachers' perceptions of principals' technology leadership is the primary purpose of this study. More specifically, this study proposes and develops a single level structural equation model to simultaneously define and measure the technology leadership construct.

## **Data Collection and Sample**

The target population consisted of 1880 teachers (i.e., tenured teachers, mentor teachers, teachers serving as administrative staffs and directors) randomly selected from 188 elementary schools within the following seven cities

in Taiwan: Keelung City, Taipei City, Hsinchu City, Taichung City, Chiayi City, Tainan City, and Kaohsiung City. These cities are located in the northern, central, and southern parts of the island respectively, and this sample is indicative of the broader teaching population of Taiwan. Participants were selected after the pilot study, and a questionnaire was sent to each teacher along with a cover letter explaining the purpose of the study and providing needed information to properly complete the questionnaire. Teachers were asked to evaluate their principal's role in leading and facilitating technology use in their schools. To encourage the return rate, follow up thank you postcards and personal contacts were made until a satisfactory percentage of participants completed the questionnaire. More importantly, teachers were told that confidentiality was assured to alleviate teacher anonymity concerns. There were 188 schools of various sizes in the sampling. The schools were grouped according to the number of classes in each school and were placed into the following categories: 17 or below classes, 18-26 classes, 27-35 classes, 36-44 classes, 45-53 classes, 54-62 classes and 63 or above classes. 1880 questionnaires were sent to randomly selected classroom teachers and administrative staff. The sample is representative of the population (of school sizes in Taiwan). Of the 1880 questionnaires distributed, 1028 questionnaires were returned for a 55 percent response rate. Table 1 displays the demographic characteristics of the respondents. Demographic variables provide a descriptive profile of those individuals who responded to the survey.

*Table 1. Number and Percent of Respondents by Demographics \* N=1028*

<b>Demographics</b>	<b>Frequency</b>	<b>Percentage</b>
<b>Gender</b>		
Male	366	35.6%
Female	618	60.1%
<b>Age</b>		
24 (or below) years old	34	3.3%
25-34 years old	396	38.5%
35-44 years old	400	38.9%
45-54 years old	169	16.4%
55 (or above) years old	24	2.3%
<b>School Size (N of class)</b>		
17 (or below) classes	114	11.1%
18-26 classes	84	8.2%
27-35 classes	99	9.6%
36-44 classes	213	20.7%
45-53 classes	167	16.2%
54-62 classes	45	4.4%
63 (or above) classes	301	29.3%
<b>Teaching Year</b>		
5 (or below) years	234	22.8%
6-10 years	205	19.9%
11-20 years	350	34.0%
21-30 years	187	18.2%
31 (or above)	42	4.1%
<b>Educational Level</b>		
Senior College	24	2.3%
BA or BS	211	20.5%
BA or BS (from teacher college or normal university)	549	53.4%
M.A. or M.Ed.	228	22.2%
Ph.D. or Ed.D.	1	0.1%
<b>Position</b>		
Mentor Teacher	370	36.0%
Specialist Teacher	128	12.5%
Administrative Head	345	33.6%
Director	165	16.1%

\* Numbers and percentages in the categories may not total 100 due to some missing data.

Using SPSS, the descriptive statistics of demographic variables were calculated. Of those teachers who responded to the demographic information on the instrument, respondent gender consisted of 366 (35.6%) males and 618 (60.1%)

females; respondent age ranged from 25 to 55, and most (400, or 38.9%) respondents were between 35 to 44 years. School size consisted of 301 (29.3%) with schools with 63 or more classes. Years of teaching experience varied, with 234 (22.8%) with less than 5 years, 205 (19.9%) with 6-10 years, and 350 (34.0%) with 11-20 years. Five hundred forty nine respondents (53.4%) held a BA from a teacher college or normal university, 228 (22.2%) held an MA or M. Ed. Three hundred seventy (36.0%) respondents identified themselves as mentor teachers, and 345 (33.6%) teachers were serving as administrative heads.

The instrument of this study, Elementary School Principals' Technology Leadership Questionnaire, was revised from The Dimensions and Implementation of the Elementary School Principals' Technology Leadership Questionnaire (Chin & Chang, 2006). It is conceptualized as four inter-related dimensions: (1) vision, planning and management (e.g., articulating a shared vision for technology use, developing a shared vision and long-range technology plan, using technology to efficiently manage administrative operations); (2) staff development and training (e.g., providing in-service training for specific skill acquisition, allocating resources for in-service training); (3) technology and infrastructure support (e.g., advocating adequate technology support, seeking out external funding sources for technology); and (4) evaluation, research, and assessment (e.g., implementing evaluation procedures for teachers' professional growth in technology, evaluating technology use in instructional programs).

Thirty-one Likert-type items (5-point scales) comprised the four dimensions that were hypothesized to define and measure principals' technology leadership. A response of "1" indicates that the principal never facilitates technology use in school, and a response of "5" indicates that the principal very often facilitates technology use in school. Based on the preliminary analysis, the four leadership dimensions showed consistency across the individual groups of assessment responses. The alpha coefficients (shown in parentheses) were calculated for each scale: vision, planning and management (.954); staff development and training (.945); technology and infrastructure support (.945); and evaluation and research (.955).

In addition to the four dimensions, the intervening construct representing principals' interpersonal and communication skills (e.g., demonstrating and maintaining positive relationships, understanding teacher needs and concerns) consisted of eight scaled leadership items and was also internally consistent across teacher groups (.966). Respondent demographics were also included on the instrument (e.g., gender, age, school size, teaching year, educational background, position, and school area). Five factors were extracted using Varimax Rotation from the original instrument, *The Dimensions and Implementation of the Elementary School Principals' Technology Leadership Questionnaire*. The variance factors were as follows: 64.652% (evaluation and research), 3.543% (vision, planning and management), 3.341% (interpersonal and communication skills), 2.859% (technology and infrastructure support), and 2.637% (staff development and training).

## Data Analysis

Empirical measurement of the dimensions and implementation of principals' technology leadership related to principals' effectiveness as perceived by their teachers in Taiwan is the intent of this study. The final structural equation model (SEM) was conducted using LISREL 8.52 (Question 1) to verify that interpersonal and communication skills are important antecedents to principals' overall effective technology leadership. SPSS was used to calculate means, standard deviations, and scale reliabilities (Cronbach's alpha). *T*-tests and analysis of variance (ANOVA) were performed to determine differences in teachers' technology leadership ratings based on teacher characteristics (Question 2). Finally, a qualitative thematic strategy was employed to analyze the practical problems that principals faced (Question 3).

## Findings and Discussion

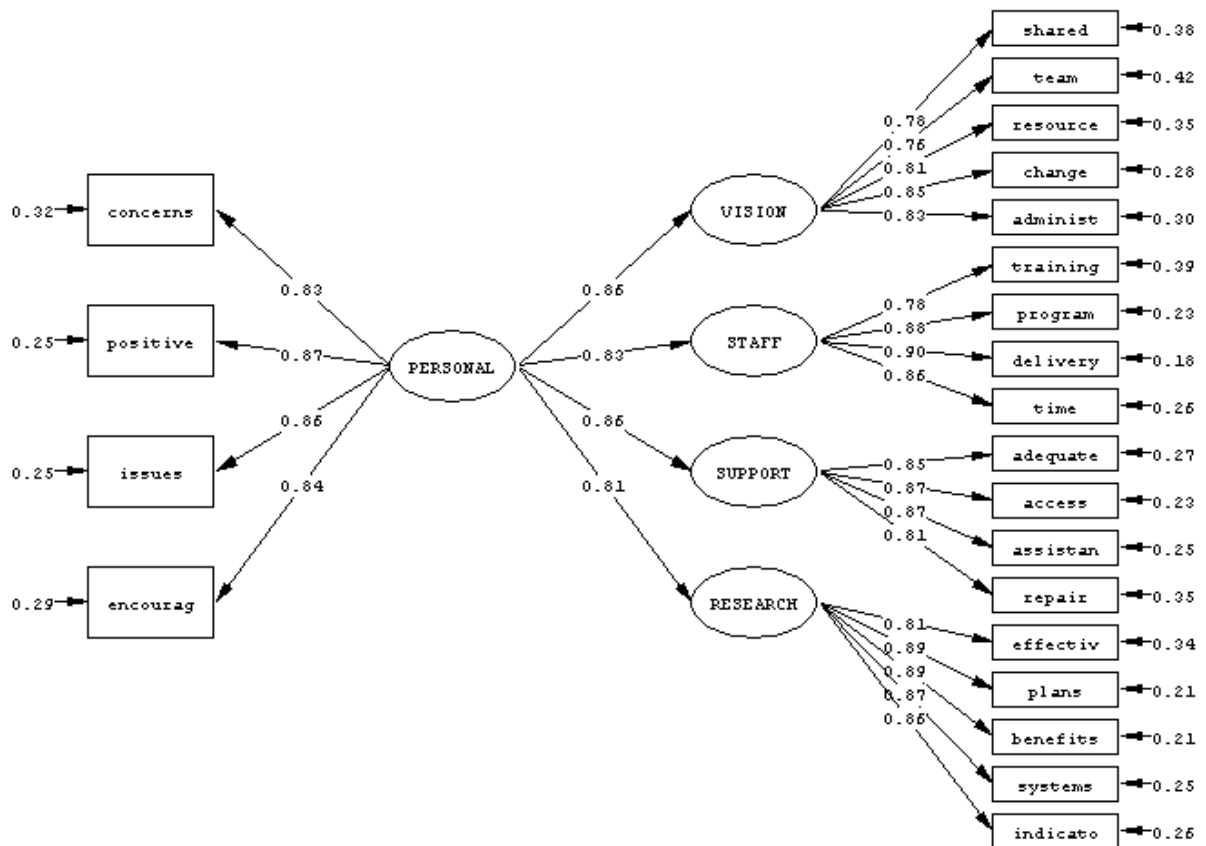
SEM provides researchers with the ability to simultaneously define and measure multidimensional constructs (e.g., principals' technology leadership). The fit of the proposed model to the data was assessed by SEM fit indices. The chi-square value is 686.44 with 201 degrees of freedom and is significant ( $p=.00$ ). The GFI, SRMR, RMSEA, NNFI, CFI, and PNFI values are .94, .034, .048, .99, .99 and .86 respectively. The Critical N is 363.14. The results of these common fit indices clearly indicate that the proposed model fit these observed data well. Table 2 displays the goodness of fit indices for the final model.



Table 2. Goodness of fit indices for the final model

Fit index	Acceptable fit	Result
Degree of Freedom=201		
Chi Square	Non-significant	686.44 (p = .01)
GFI	> .90	.94
SRMR	> .05	.034
RMSEA	> .05	.048
NNFI	> .90	.99
CFI	> .90	.99
PNFI	> .50	.86
Critical N	> 200	363.14

After ensuring that the proposed model fit the data, standardized parameter estimates in the model were considered. The four constructs (i.e., vision, staff, support, research) comprising principals' technology leadership define and measure the proposed model well, and all parameter estimates were significant (.86, .83, .86, .81, respectively) as shown in Figure 1. The results suggest that the observed data and the proposed model fit these data quite well; that is, the dimensions defined effective technology leadership well. In other words, vision, planning, and management; staff development and training; technology and infrastructure support; and evaluation and research are the four dimensions of behavior that explain the effective technology leadership of principals.



Chi-Square=686.44, df=201, P-value=0.00000, RMSEA=0.048

Figure 1. The model of principals' technology leadership

These results indicate that principals need to develop and implement a visionary long-range technology plan in order to be effective technology leaders. The findings also show the importance to principals of staff development and training activities for their teachers and students. Principals must also ensure that the school's technology infrastructure is well supported, and as leaders of technology they must develop an evaluation and assessment plan for their schools. These dimensions of behavior significantly explain principals' effective technology leadership.

Principals' interpersonal and communication skills showed a significant and positive impact on teachers' perceptions of principals' effective technology leadership. This finding supports previous research (Chang, 2002) showing that to become effective technology leaders, principals must build positive working relationships, communicate change and new ideas well, and identify and support teacher needs and concerns. As technology leaders, principals who embrace technology will effectively lead their schools to acquire educational resources to enhance student engagement and learning. Drawn from Figure 1, the dimensions and performance indicators of principals' technology leadership are shown in Table 3.

*Table 3. The Dimensions and Performance Indicators of Principals' Technology Leadership*

<b>Dimension</b>	<b>Performance Indicator</b>
Vision, Planning and Management	<ol style="list-style-type: none"> <li>1. Clearly articulates a shared vision for technology use in the school</li> <li>2. Empowers a diverse and inclusive technology planning team</li> <li>3. Advocates for school technology resources</li> <li>4. Manages technology change effectively</li> <li>5. Uses technology to efficiently manage administrative operations</li> </ol>
Staff Development and Training	<ol style="list-style-type: none"> <li>1. Encourages technology in-service training</li> <li>2. Supports technology in-service training program design</li> <li>3. Supports technology in-service training delivery</li> <li>4. Provides technology training release time</li> </ol>
Technology and Infrastructure Support	<ol style="list-style-type: none"> <li>1. Ensures appropriate technology facilities</li> <li>2. Ensures equal access to technology resources</li> <li>3. Ensures technology support to school personnel when assistance is needed</li> <li>4. Ensures equipment timely repair and maintenance</li> </ol>
Evaluation and Research	<ol style="list-style-type: none"> <li>1. Considers effective technology use as one performance assessment component of instructional staff</li> <li>2. Evaluates school technology plans</li> <li>3. Evaluates technology in terms of costs/benefits</li> <li>4. Evaluates computer operating systems for classrooms and laboratories</li> <li>5. Utilizes district level data to evaluate technology instructional use</li> </ol>
Interpersonal and Communication Skills	<ol style="list-style-type: none"> <li>1. Demonstrates an understanding of technology needs and concerns of faculty, staff and students</li> <li>2. Maintains positive relationships with faculty, staff and students in regard to technology</li> <li>3. Communicates effectively with faculty, staff, and students about technology</li> <li>4. Encourages school personnel to utilize information sources about technology for professional development</li> </ol>

Teacher demographic characteristics (e.g., gender, age, school size, teaching year, position and school area) and ratings of the four-technology leadership dimensions (e.g., vision, staff, support, and research) were significantly related at the .05 alpha level, as shown by t-test and ANOVA results. Table 4 displays the mean and standard deviation of the dimensions of technology leadership. Among the five dimensions analyzed, the mean of "*Interpersonal and Communication Skills*" (Mean=3.929, SD=.736) was highest, with the other dimensions as follows: *Technology and Infrastructure Support* (Mean=3.913, SD=.747), *Staff Development and Training* (Mean=3.775, SD=.758), *Vision, Planning and Management* (Mean=3.758, SD=.724) and *Evaluation and Research* (Mean=3.521, SD=.787). Teachers' perceptions of principals' technology leadership implementation were positive, based on overall analysis.

*Table 4. Mean and standard deviation of the dimensions of technology leadership*

<b>Dimension</b>	<b>Mean</b>	<b>Standard Deviation</b>
Interpersonal and Communication Skills	3.929	0.736
Technology and Infrastructure Support	3.913	0.747
Staff Development and Training	3.775	0.758
Vision, Planning and Management	3.758	0.724
Evaluation and Research	3.521	0.787

Demographic and environmental variables, including age, school size, teaching years, educational level, position, and school area were significant (t test and ANOVA) except for the educational level variable. Tables 5 to 9 display the analysis of variance for demographic and environmental variables on the dimensions of technology leadership.

*Table 5. Analysis of Variance for Age on the Dimensions of Technology Leadership*

		<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
Vision	Between Groups	10.155	4	2.539	4.976	.001
	Within Groups	495.874	972	.510		
	Total	506.029	976			
Training	Between Groups	15.008	4	3.752	6.865	.000
	Within Groups	537.235	983	.547		
	Total	552.242	987			
Support	Between Groups	9.251	4	2.313	4.375	.002
	Within Groups	516.430	977	.529		
	Total	525.680	981			
Evaluation	Between Groups	10.980	4	2.745	4.602	.001
	Within Groups	581.518	975	.596		
	Total	592.498	979			
Skill	Between Groups	12.844	4	3.211	6.106	.000
	Within Groups	517.441	984	.526		
	Total	530.285	988			

*Table 6. Analysis of Variance for School Size on the Dimensions of Technology Leadership*

		<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
Vision	Between Groups	7.530	6	1.255	2.446	.024
	Within Groups	497.272	969	.513		
	Total	504.802	975			
Support	Between Groups	9.208	6	1.535	2.900	.008
	Within Groups	516.054	975	.529		
	Total	525.262	981			
Evaluation	Between Groups	7.931	6	1.322	2.203	.041
	Within Groups	583.859	973	.600		
	Total	591.790	979			
Skill	Between Groups	10.573	6	1.762	3.342	.003
	Within Groups	517.805	982	.527		
	Total	528.378	988			

*Table 7. Analysis of Variance for Teaching Year on the Dimensions of Technology Leadership*

		<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
Vision	Between Groups	13.510	4	3.377	6.667	.000
	Within Groups	489.344	966	.507		
	Total	502.854	970			
Training	Between Groups	19.633	4	4.908	9.067	.000
	Within Groups	529.410	978	.541		
	Total	549.042	982			
Support	Between Groups	7.254	4	1.814	3.410	.009
	Within Groups	516.934	972	.532		
	Total	524.188	976			
Evaluation	Between Groups	11.251	4	2.813	4.710	.001
	Within Groups	579.316	970	.597		
	Total	590.567	974			
Skill	Between Groups	15.446	4	3.861	7.384	.000
	Within Groups	511.986	979	.523		
	Total	527.432	983			

Table 8. Analysis of Variance for Position on the Dimensions of Technology Leadership

		Sum of Squares	df	Mean Square	F	Sig.
Vision	Between Groups	8.625	3	2.875	5.624	.001
	Within Groups	490.792	960	.511		
	Total	499.417	963			
Training	Between Groups	18.371	3	6.124	11.290	.000
	Within Groups	527.222	972	.542		
	Total	545.593	975			
Support	Between Groups	15.069	3	5.023	9.613	.000
	Within Groups	504.742	966	.523		
	Total	519.811	969			
Evaluation	Between Groups	8.561	3	2.854	4.784	.003
	Within Groups	575.029	964	.597		
	Total	583.590	967			
Skill	Between Groups	12.418	3	4.139	7.868	.000
	Within Groups	512.431	974	.526		
	Total	524.849	977			

Table 9. Analysis of Variance for School Area on the Dimensions of Technology Leadership

		Sum of Squares	df	Mean Square	F	Sig.
Vision	Between Groups	11.750	3	3.917	7.690	.000
	Within Groups	484.877	952	.509		
	Total	496.628	955			
Training	Between Groups	7.039	3	2.346	4.205	.006
	Within Groups	537.940	964	.558		
	Total	544.978	967			
Support	Between Groups	9.107	3	3.036	5.775	.001
	Within Groups	504.124	959	.526		
	Total	513.231	962			
Evaluation	Between Groups	15.979	3	5.326	9.092	.000
	Within Groups	561.801	959	.586		
	Total	577.779	962			
Skill	Between Groups	5.685	3	1.895	3.580	.014
	Within Groups	510.840	965	.529		
	Total	516.525	968			

Each study participant in the questionnaire was also asked, “*What’s your opinion regarding your principal’s role in facilitating technology use in your school?*” Using open and axial coding techniques, participants’ answers (transcripts) were thematically aggregated for important emerging category analysis. Four themes emerged from the transcript data: budget shortage problems, technology facility problems, staff development problems, and leadership problems. Schools are frequently faced with budget shortages. Teachers recommended that their principals raise technology funds. For example, teachers replied:

Our classrooms should have Internet connections and computer equipment set ups; integrating instructional technology is the intent, but insufficient funds result in insufficient equipment; therefore principals should raise funds for computer upgrades.

Participants also recommended that principals seek out external funding sources to provide applicable technology equipment. For example:

Principals should seek out external support and technology funding resources in Order to upgrade technology equipment and provide adequate technology to school members.

Expenditure was associated with the problems principals encountered with technology facilities. Expense directly influenced the availability of adequate facilities, which impacted technology use. Principals, according to respondents, should provide instructional equipment, hardware and software to meet faculty and student needs. For example, participants responded:

It is necessary to evaluate classroom projector needs; hardware and software should be regularly upgraded, and faculty and students should have unlimited access to facilities; insufficient instructional facilities and dysfunctional multimedia should be improved; computer and peripheral equipment (e.g., Internet connection, printer, instructional software) should be provided for each grade.

Teachers also indicated that schools should ensure that technology facilities were appropriate for instruction and learning. For example:

Teachers could not integrate technology into classroom instruction due to insufficient equipment.  
Schools should remodel infrastructures to facilitate teacher's professional technology development.

One participant, talking about his school, positively commented:

Two laser pointers, a printer, and an ink cartridge were provided to each teacher for convenience in utilizing audio and video equipment.

Staff development problems encountered by principals were associated with teachers' technology literacy, which was correlated with in-service training. Participants replied that schools should provide technology workshops for in-service teachers. For example:

In order to promote teachers' technology literacy, technology workshops should be offered; principals should play an advocacy role and hold a series of school-based technology workshops to facilitate technology use.

Teachers also pointed out cases where their schools promoted teachers' technology literacy. For example:

I appreciated my principal's support in helping teachers integrate technology into instruction, creating websites, and handling student affairs; our principal prioritized computer instruction as the key feature of students' curriculum, and our building had sufficient equipment. The principal also demonstrated an understanding of faculty, staff, and student technology needs and concerns.

Principals' leadership responsibilities are to lead and support. For example, teachers stated that:

Principals should implement their leadership roles and strengthen their technology capacity; our principals advocated adequate technology support, and encouraged faculty and students to use technology.

One participant expressed that principals do not necessary need to be technology specialists. For example:

A principal must empower his team members to manage technology, seek out external technology resources, and encourage faculty technology use instead of becoming a technology specialist.

Based on the aforementioned findings, elementary principal technology leadership has been defined and measured as one construct comprising four technology leadership dimensions (i.e., vision, planning, and management; staff development and training; technology and infrastructure support; evaluation and research). The first dimension, vision, planning, and management, implies that, as technology leaders, principals should clearly articulate a shared technology vision for their schools. Previous research shows that articulating, sharing, and demonstrating a technology vision for the schools are necessary effective leadership behaviors (Aten, 1996; Cory, 1990; Ford, 2000; Inkster, 1998; Jewell, 1998; Ray, 1992). Furthermore, organizing and empowering a technology-planning team are important vision and planning responsibilities for teachers' technology leadership perceptions. Managing technology resources effectively is also an essential for principals. This study supports previous research in that vision, planning, and managing technology resources are critical elements to principal overall technology leadership. Therefore, effective principals should advocate for school technology resources, exercise fair and reasonable judgment in allocating technology funds and resources, implement a reasonable technology plan, and manage technology facilities and resources well.

The second dimension, staff development and training, suggests that principals should help train and encourage teacher's technology development. Staff technology development and training is an essential aspect of principals' technology leadership. To demonstrate technology leadership effectively, principals should encourage teachers' technology skill acquisition, plan and design on-going and future technology staff development programs, and provide teachers adequate time for technology training. Staff development and training significantly explained principal technology leadership in this case. These findings also support Spence's (1999) (as cited in Ford, 2000)

study examining principals' technology leadership competencies. He argued that identifying and coordinating appropriate staff development activities was listed as the single most important technology leader responsibility. Similarly, Aten's (1996) study showed that staff development was the number one skill most needed. Teachers highly value technical assistance in developing and nurturing technology skills. This study also supports previous research that staff development and training is an important principal technology leadership characteristic.

The third dimension, technology and infrastructure support, argues that principals need to provide adequate technology support. Acquiring technology, as well as maintaining and supporting school infrastructure, are crucial areas of principals' technology leadership. Principals, as technology leaders, must ensure appropriate technology facilities, provide access to technology resources, and support school personnel when technical assistance is needed. Principals must ensure timely repair and maintenance of equipment in their schools. Technology dimension and infrastructure support significantly explained principal's effective technology leadership in this study. Simply put, advocating adequate technology and infrastructure support for school members (e.g., teachers, staff members, and students) are important principal technology leadership characteristics.

The fourth dimension, evaluation and research, suggests that principals as technology leaders should consider technology use as one component in instructional staff performance assessment. They also need to evaluate school technology plans in terms of costs/benefits, and monitor computer operating systems in both classrooms and laboratories. More importantly, principals should also utilize district level data to evaluate instructional technology use. Furthermore, comparing school data with district and national data can often provide ideas, trends, and successes [failures] applicable to school performance and effectiveness improvement. Evaluation and research significantly contributed to overall principal technology leadership in this study. This dimension reflects previous research showing important principal technology leadership characteristics. Principals as technology leaders also need strong interpersonal and communication skills to be effective. They should demonstrate an understanding of technology needs and concerns of teachers, staff members, and students. More important, it is imperative that principals maintain positive and constructive interpersonal relationships, and communicate effectively with their teachers, staff and students, and encourage school personnel to utilize technology information sources for professional growth.

## **Conclusions and Implications**

Researchers have recently turned to the study of technology leadership in terms of technology's support of school reform. Application of leadership skills necessary for school leaders to help their institutions apply technology in beneficial ways and prepare their schools for the 21<sup>st</sup> century is the meaning of technology leadership. Researchers have stated that building principals' technology leadership is essential in schools; principals must model effective technology leadership. In this study, effective technology leadership of principals has been defined and measured as one construct comprising four dimensions (domains). The method used in this study, Structural Equation Modeling (SEM), offers an advancement in the ability to define the multidimensional technology leadership construct. In this case, effective technology leadership of principals was quite well defined and well measured. A technology leader is one who leads the school in improvement on restructuring, and uses emerging technologies as the core resources for educational change. More importantly, the role of principal is now evolving to that of a technology leader. To be an effective technology leader, a principal should develop and implement a technology vision and long-range technology plan in the school. A principal requires a sense of vision, since technology leadership is the ability to develop and articulate a vision of how technology can produce change. Moreover, a principal should encourage faculty development in technology. Principals should plan and design staff development activities for their school settings. Beyond that, advocating technological support, ensuring that facilities for technology use are adequate, and evaluating school and district technology plans are the roles and responsibilities of the principal as technology leader.

The paper is significant since numerous technology literacy-training initiatives have been implemented for teachers in Taiwan during the past decade. With relatively few studies specifically addressing evaluation of building principals' technology leadership in Taiwan, this is an area that necessitates future exploration so that current and future leaders can be prepared to deal more effectively with technology and can be expected to successfully implement technology policy. The evaluative instrument and process for assessing technology leadership discussed has merit. More importantly, the paper extends work in the technology field, and the paper offers the alternative

points of view that interpersonal and communication skills are important antecedents to principals' overall effective technology leadership. The aforementioned finding makes a significant contribution to the field.

### **Implications for Theory**

Although there is an abundance of studies related to technology leadership in the educational literature, few studies have focused on measuring principals' effective technology leadership as perceived by their teachers. The findings from this study support the idea that effective technology leadership of principals can be empirically defined and measured. The effective technology leadership constructs reported in this study both support and add to previous research findings.

### **Implications for Single Level Analysis in Leadership Studies**

Previous studies on technology leadership focused on descriptive statistical analysis. A challenge to this study was employing the technique of Structural Equation Modeling. Structural Equation Modeling offers advancement in the ability to simultaneously define multidimensional constructs such as technology leadership and also to test for direct and indirect effects of technology leadership performance on principals' effectiveness. As proposed in this study, an SEM covariance structure analysis can be extended to include sets of predictors at the individual level. The results of this study show that SEM is a powerful means of defining and measuring multidimensional constructs of technology leadership.

### **Implications for Practice and Policy**

Much current research notes that the principal's role has evolved to that of an effective technology leader. Moreover, researchers have shown that principals' technology leadership is essential in schools. As this study indicates, principals as technology leaders must develop and implement a school vision and technology plan, encourage teacher technology development and training, provide adequate infrastructure and technology support, and develop an effective school evaluation plan. Principals who embrace their evolving role as technology leaders can effectively lead and prepare their schools. More than ever, principals must lead their schools in acquiring and using new and emerging technologies as educational resources for enhanced student engagement and learning. The four technology leadership dimensions examined in this study provide principals with the knowledge and skills necessary to use technology and enhance their effectiveness as school leaders. Moreover, this evaluation instrument may provide educational administrators with substantive information to enhance the recruitment and selection of principals as technology leaders. As a result of this evaluation and assessment information, Taiwan's Department of Education could sponsor preparation programs providing professional development for principals to improve classroom technology use, evaluate teacher and student strengths and needs in technology, and develop a practical and useful technology plan. Increased student learning and achievement through the application of new and emerging educational technologies is the primary goal.

### **References**

- Anderson, R. E., & Dexter, S. L. (2000). *School technology leadership: Incidence and impact. Teaching, learning, and computing: 1998 National survey, report #6*, ERIC Document Reproduction Service No. ED449786.
- Anderson, R. E., & Dexter, S. (2005). School technology leadership: An empirical investigation of prevalence and effect. *Educational Administration Quarterly*, 41(1), 49-82.
- Appalachia Educational Lab (2000). *Principal connections: A guide to technology leadership*, ERIC Document Reproduction Service No. ED443384.
- Aten, B. M. (1996). *An analysis of the nature of educational technology leadership in California's SB 1274 restructuring schools*, Unpublished doctoral dissertation, University of San Francisco.
- Bailey, G. D., & Lumley, D. (1994). *Technology staff development programs. A leadership sourcebook for school administrators*, New York: Scholastic.

- Bailey, G. D. (1997). What technology leaders need to know: The essential top 10 concepts for technology integration in the 21<sup>st</sup> century? *Learning & Leading with Technology*, 25(1), 57-62.
- Battle, M. C. (2004). *Superintendents' perceptions of educational technology leadership*, Unpublished doctoral dissertation, The University of Southern Mississippi.
- Bridges, J. W. (2003). Principal influence: Sustaining a vision for powerful new forms of learning using technology (Doctoral dissertation, University of California, Los Angeles). *Dissertation Abstracts International*, 64(6), 1915.
- Brush, T. (1998). *Technology planning and implementation in public schools: A five-state comparison*, ERIC Document Reproduction Service No. ED419528.
- Calhoun, K. J. (2004). *Superintendent change leadership strategies associated with successful technology integration in public school districts*, Unpublished doctoral dissertation, University of Laverne.
- Centre for Information Technology in Education (2005). *IT leadership training course for primary school principals*, retrieved June 10, 2008 from <http://eleadership.cite.hku.hk>.
- Chang, I. (2002). *Assessing principals' leadership in implementing educational technology policies: An application of structural equation modeling*, Unpublished doctoral dissertation, University of Missouri-Columbia.
- Chang, I. (2003a). The new leadership trend: Technology leadership. *Journal of Education Research*, 114, 83-95.
- Chang, I. (2003b). Assessing the dimensions of principals' effective technology leadership: An application of structural equation modeling. *Educational Policy Forum*, 6(1), 111-141.
- Chang, I. (2004). *A study of the elementary school principals' technology leadership effectiveness*, National Science Council (93-2413-H-260-007), Taiwan.
- Chang, I., & Tseng, D. (2005). The development of technology leadership academies in the USA and from which some implications for school administration in Taiwan. *Journal of the National Institute for Complication and Translation*, 33(3), 83-95.
- Chang, M. (2005). *The leadership ideas and strategies of learning schools*, retrieved June 13, 2008 from <http://www.edpl.tku.edu.tw/2-04.doc>.
- Checkley, K. (2000). The contemporary principal: New skills for a new age. *Education Update*, 42(3), 1-8.
- Cheng, Y. (2004). *A study of relationship between Miaoli County's elementary school principals' instructional leadership and school effectiveness*, Unpublished master's thesis, National Taichung Teachers College, Taiwan.
- Chin, J. M., & Chang, I. (2006). The study of the dimensions and implementation of the elementary school principals' technology leadership. *Journal of Education & Psychology*, 29(1), 1-27.
- Collis, B. (1988). *Computers, curriculum, and whole-class instruction*, Belmont: Wadsworth.
- Cory, S. (1990). Can your district become an instructional technology leader? *The School Administrator, Special Issue*, 17-19.
- ED Office of Educational Technology (2001). *E-learning: Putting a world-class education at the fingertips of all children*, retrieved June 13, 2008 from <http://www.ed.gov/Technology/elearning/index.html>.
- Ferris, A., & Roberts, N. (1994). Teachers as technology leaders: Five case studies. In Kearsley, G. & Lynch, W. (Eds), *Educational technology: Leadership perspectives*, New Jersey: Educational Technology Publications, 53-74.
- Ford, J. I. (2000). *Identifying technology leadership competencies for Nebraska's K-12 technology leaders*, Unpublished doctoral dissertation, University of Nebraska, USA.
- Frazier, M. K. (2003). *The technology coordinator in K-12 school districts: The research, development, and validation of a technology leader's guide*, Unpublished doctoral dissertation, Kansas State University, USA.
- Glatthorn, A. A. (2000). *The principal as curriculum leader: Shaping what is taught and tested*, Thousand Oaks: Corwin Press.
- Glennan, T. K., & Melmed, A. (1996). *Fostering the use of educational technology: Elements of a national strategy*, Santa Monica: RAND.
- Hsieh, W. B. (2004). The new dimension of professional development of principals elementary and secondary schools: Technology leadership. *Secondary Education*, 55(5), 72-91.



- Hudanich, N. V. (2002). *Identifying educational technology leadership competencies for New Jersey's school superintendents*. Unpublished doctoral dissertation, Seton Hall University, USA.
- Hughes, M., & Zachariah, S. (2001). *An investigation into the relationship between effective administrative leadership styles and the use of technology*, retrieved July 20, 2008 from <http://www.ucalgary.ca/~iejll/volume5/hughes.html>.
- Huang, H. (2004). *A study of elementary school principals' instructional leadership strategy in Taipei County*, Unpublished master's thesis, National Taiwan Normal University, Taiwan.
- Inkster, C. D. (1998). *Technology leadership in elementary school principals: A comparative case study*, Unpublished doctoral dissertation, University of Minnesota, USA.
- ISTE (1998). *Curriculum guidelines for accreditation of educational computing and technology progress: A polio preparation manual*, Eugene: International Society for Technology in Education.
- ISTE (2001). *Role-specific technology leadership tasks: Principal DRAFT v4.0*, retrieved July 14, 2008 from <http://cnets.iste.org/tssa/printtaskprofile.html>.
- ISTE (2005). *Educational technology standards and performance indicators for administrators*, retrieved July 28, 2008 from [http://cnets.iste.org/administrators/a\\_stands.html](http://cnets.iste.org/administrators/a_stands.html).
- Jewell, M. J. (1998). The art and craft of technology leadership. *Learning and Leading with Technology*, 26(4), 46-47.
- Kadela, T. (2002). *Technology leadership of elementary principals: Standards, competencies, and integration*, Unpublished doctoral dissertation, Seton Hall University, USA.
- Kearsley, G., & Lynch, W. (1994). Educational leadership in the age of technology: The new skills. In Kearsley, G. & Lynch, W. (Eds.), *Educational technology: Leadership perspectives*, New Jersey: Educational Technology, 5-17.
- Keating, D. A., Stanford, T. G., Self, A., & Monniot, J. (1999). Graduate professional education of engineers in industry for innovation and technology leadership. *Industry & Higher Education*, 13(1), 33-39.
- Kinnaman, D. E. (1994). What it really means to integrate technology. *Technology & Learning*, 14(8), 130.
- Kline, J. J. (1993). *A comparative analysis of selected educational technology competencies regarded as important for the integration of technology in the public schools of Pennsylvania*, Unpublished doctoral dissertation, Lehigh University, USA.
- Kowch, E., & Walker, K. (1996). Toulmin's moral reasoning model applied to ethical Internet choices: A means for exercising ethical technology leadership. *Canadian Journal of Educational Communication*, 25(3), 235-54.
- Kozloski, K. C. (2006). *Principal leadership for technology integration: A study of principal technology leadership*, Unpublished doctoral dissertation, Drexel University.
- MacNeil, A. J., & Delafield, D. P. (1998). *Principal leadership for successful school technology implementation*, ERIC Document Reproduction Service No. ED421126.
- Martens, J. D. (2003). *Shooting star leadership: Identifying the key attributes for IT leaders during chaotic times*, Unpublished doctoral dissertation, Colorado Technical University, USA.
- Matthews, A. W. (2002). *Technology leadership at a junior high school: A qualitative case study*, Unpublished doctoral dissertation, University of Nevada-Las Vegas, USA.
- Mirra, D. R. (2004). *The role of the school superintendent as a technology leader: A Delphi study*, Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, USA.
- Moursund, D. (1992). *The technology coordinator*, Eugene: International Society for Technology in Education.
- Murphy, D. T., & Gunter, G. A. (1997). Technology integration: The importance of administrative supports. *Educational Media International*, 34(3), 136-139.
- Nash, G. W. (2002). *A case study of the superintendent as technology leader in an East Texas school district*, Unpublished doctoral dissertation, Stephen F. Austin State University, USA.
- No Child Left Behind (2001). *Enhancing education through technology*, retrieved June 18, 2008 from <http://www.ed.gov/inits/nclb/partx.html>.

- Orr, M. T., & Barber, M. E. (2006). Collaborative leadership preparation: A comparative study of partnership and conventional programs and practices. *Journal of School Leadership, 16*(6), 703-739.
- Ray, D. (1992). Educational technology leadership for the age of restructuring. *The Computing Teacher, 19*(6), 8-14.
- Reeves, D. B. (2004). Evaluating administrators. *Educational Leadership, 61*(7), 52-58.
- Robinson, B. (1994). Technology leadership in the English educational system: From computer systems to systematic management of computers. In Kearsley, G. & Lynch, W. (Eds.), *Educational technology: Leadership perspectives*, New Jersey: Educational Technology, 137-152.
- Rogers, B. A. (2000). *The correlation between teachers' perceptions of principals' technology leadership and the integration of educational technology*, Unpublished doctoral dissertation, Ball State University, USA.
- Ross, T. W. (1993). Research, development, and validation of a principal's handbook for implementing technology-based learning methods in information-age school (Doctoral dissertation, Kansas State University, 1993). *Dissertation Abstracts International, 54*(8), 2835.
- Ross, T. W., & Bailey, G. D. (1996). *Technology-based learning: A handbook for teachers and technology leaders*, Arlington Heights: IRI/Skylight.
- Scanga, D. (2004). *Technology competencies for school administrators: Development and validation study of a self-assessment instrument*, Unpublished doctoral dissertation, University of South Florida, USA.
- Scott, G. (2005). *Educator perceptions of principal technology leadership competencies*, Unpublished doctoral dissertation, The University of Oklahoma, USA.
- Seay, D. A. (2004). *A study of the technology leadership of Texas high school principals*, Unpublished doctoral dissertation, University of North Texas, USA.
- Sharp, W., & Walter, J. K. (1994). *The principal as school manager*, Lancaster: Technomic.
- Shen, C. (2004). *Technology and learning: Theory and practice*, Taipei: Psychology Press.
- Shulman, M. D. (2003). *Superintendent conceptions of the institutional conditions that facilitate or impede the process of teacher technology integration: A qualitative study of three New Hampshire school superintendents*, Unpublished doctoral dissertation, University of Massachusetts Lowell, USA.
- Stegall, P. (1998). *The principal: Key to technology implementation*, ERIC Document Reproduction Service No. ED424614.
- Thomas, L. G., & Knezek, D. (1991). Providing technology leadership for restructured schools. *Journal of Research on Computing in Education, 24*(2), 265-279.
- U.S. Department of Education (2001). *A nation at risk*, retrieved August 30, 2007 from <http://www.ed.gov/pubs/NatAtRisk/index.html>.
- Ury, G. G. (2003). *Missouri public school principals' computer usage and conformity to technology standards*, Unpublished doctoral dissertation, University of Missouri-Columbia, USA.
- Wagner, W. W. (2004). *The technology coordinator: Key characteristics and traits of successful educational technology leaders*, Unpublished doctoral dissertation, Ashland University, USA.
- Weber, M. J. (2006). *A study of computer technology use and technology leadership of Texas elementary public school principals*, Unpublished doctoral dissertation, University of North Texas, USA.
- White, D. E. (2004). *A study of relationships between selected demographic variables and the principals' self-perception of technology leadership*, Unpublished doctoral dissertation, Illinois State University, USA.
- Wexler, S. (1996). *Bill Gates goes back to school*, retrieved August 19, 2008 from <http://cdmom.vergant.com/hitech/edu.html>.
- Wu, H. L. (2004). *A study of elementary school principals' curriculum leadership in Taipei City*, Unpublished master's thesis, National Chungcheng University, Taiwan.
- Yeh, L. (2003). Technology leadership. *Journal of Education Research, 112*, 151-152.
- Yoho, J. K. (2006). *Technology leadership, technology integration, and school performance in reading and math: A correlation study in K-12 public schools*, Unpublished doctoral dissertation, Pennsylvania State University, USA.

## The Impact of Externalizing Readers' Mental Representation on the Comprehension of Online Texts

Yu-Fen Yang and Ya-Chen Hung

Graduate School of Applied Foreign Languages, National Yunlin University of Science & Technology, Taiwan // Tel: 886-5-534-2601 Ext. 3136 // Fax: 886-5-531-2058 // yangy@yuntech.edu.tw

### ABSTRACT

This study reports on our design of a computer system which supports the understanding of how students construct their mental representation of references and how the construction failures impede their reading comprehension. Three modules, *User interface*, *Recording*, and *Feedback* were implemented. The *recording module* traced all students' reading process when they tried to link sentences together by references. Results showed that more-proficient readers' mental process in resolving references was constructed to be a complete and coherent network. When they encountered reading difficulty, they usually engaged in comprehension monitoring to read and reread the related sentences to find out contextual clues. They also asked for a *feedback tool* for assistance. All these helped them grasp the main idea of a text and solve lexical ambiguities. In contrast, average and less-proficient readers often resolved the references separately or mismatched them to an incorrect subject which led to partial understanding of textual information. Their unsuccessful resolution of references in previous sentences often hindered their interpretation of subsequent sentences. This caused repeated reading failures in the comprehension of an online text. The mental maps and reading process shown in this study clearly explain students' reading success and failure. The discussion of these maps and processes between the teacher and students will be one of the ways to promote latter's awareness in employing comprehension monitoring and referring strategy.

### Keywords

Mental representation, Referential resolution, Reading failure, Process data, and Main idea

### Introduction

In the information age, reading enormous information resources provided by internet underpin professional success. That is why the development of information processing abilities is one of the primary objectives in higher education; particularly, college reading is considered an information-processing skill that a student should internalize for meaning construction (Yang, 2007). In college, reading is not only essential to academic learning in all content areas but also to professional success and, indeed, to lifelong learning (Levine, Ferenz, & Reves, 2000; Pritchard, Romeo, Muller, 1999). Nevertheless, many students who register for undergraduate study are under-prepared for university education and have low levels of reading proficiency (Dreyer & Nel, 2003). This is particularly true for college students who learn English as a Foreign Language (EFL) in Taiwan.

Many EFL college students in Taiwan were found to lack reading skills necessary for success in highly demanding college courses (Huang, 2005; Yang, 2003, 2004). They often encountered difficulties and failures in the process of reading comprehension. According to Chen, You, Yang, & Huang's study (2004), 35.5% of them failed the reading section at the elementary level of *General English Proficiency Test*, a nationwide screening test. That is, many of them did not have the required ability for college reading, since their reading proficiency was identified as being equivalent to junior high school level.

When pressed to read, EFL college students often selected ineffective and inefficient strategies with little strategic intent (Wood, Motz, and Willoughby, 1998). Often this was due to their low level of reading strategy knowledge and lack of comprehension monitoring engagement (Dreyer, 1998). Another factor might be their inexperience coming from the limited task demands of high school and the focus of traditional reading instruction on the teacher's one-way lecture and knowledge reproduction (Dreyer and Nel, 2003). As a result, some students would not engage in reading strategy or comprehension monitoring unless they were asked to think about their reading process through activities or instruction (Hartley, 2001).

The computer assisted learning environment was shown to greatly support the reader's engagement of reading strategy and comprehension monitoring for it provided explicit modeling and individualized scaffolding (Potelle & Rouet, 2003). The modeling and scaffolding assisted the reader to manage and monitor his/her own process in

reading various texts. This is fundamentally important for EFL college readers in Taiwan for there are about 45 or more students of varying language proficiency levels involved in one class. The large class size limits the classroom teacher to provide the individualized support, guidance, and monitoring required for each reader's progress.

To monitor readers' progress, Chang, Sung, and Chen (2002) investigated the effects of concept-mapping strategies in a computer system. They designed three approaches—map correction, scaffold fading, and map generation—to enhance students' text comprehension and summarization abilities. The experimental results showed that the map-correction method enhanced text comprehension and summarization abilities and that the scaffold-fading method facilitated summarization ability. Chang et al. addressed that graphic strategies, such as graphic organizers and knowledge maps, had proved useful for text comprehension. In addition, Vakilifard, Armand, and Baron (2006) examined whether concept maps enhanced the second language students' reading comprehension in the first language context. Their results showed that the early presentation of concept maps before reading texts led students in experimental group to understand the texts effectively. Positive results of using concept maps in various subject areas could be found in Hazzan's (2004), Kwon and Cifuentes' (2007), Ruiz-Primo, Schultz, Li, and Shavelson's (2001) studies.

In the literature reviewed, very few studies addressed college students' mental representation of reading process (Yang, Yeh, Wong, and Lee, 2008). The process data were rarely found to indicate how students read and reread a sentence, how they selected, deleted, and reselected a text element, and how they integrated different sentences by building connections between text elements. These important data were even hardly reported back to the teacher and the individual student (Schacter, Herl, Chung, Dennis, and O'Neil, 1999). The teacher got no chance to monitor his/her students' reading process in details. Similarly, the individual student lost the opportunity to reflect on their own process. This study reports on our design of a computer system that supports the understanding of how students construct their mental representation of references and how the construction failures impede their reading comprehension. In the system, each student's reading behavior and process were recorded by a *recording module*. It made the intangible reading process visible to the teacher. This also enabled the teacher to observe the difficulties that his/her students encountered and the differences in performance among students with various English reading proficiencies. Based on this information, the teacher was able to modify his/her follow-up instruction to help students overcome their difficulties and better develop their processing skills in reading.

In this study, mental representation is defined as the mental map that demonstrates the reader's cognitive structures for establishing the relationships among the references in different parts of a text. In the mental map, references are represented as nodes and connected by lines. The connecting lines express the relationships among the references and make the mental map more meaningful by showing how the student relates references to each other in his/her cognitive structure. This is also called co-reference or referential resolution (Walsh & John-Laird, 2004). Referential resolution is a reading strategy applied by the reader to interpret the references that have the same meanings as other elements in the text, such as "he" refers to "John." While resolving the references, the reader is actually engaged in comprehension monitoring which he/she monitors, regulates, and evaluates his/her own reading process (Hartman, 2001). The management and regulation of one's own reading process is helpful for meaning construction of text (Paris and Winograd, 1990).

## Method

### Participants

A total of 92 junior and senior college students were recruited from three reading classes in a technological university in central Taiwan. Their language proficiency levels were identified by their reading scores in a simulated online exam *Testing of English for International Communication* (TOEIC) at <http://140.125.169.80>. The maximum score in the reading section of the online exam was 200.

The frequency distribution of all the participants' score was used to divide the participants into three groups of readers. The frequency distribution is shown in Figure 1.

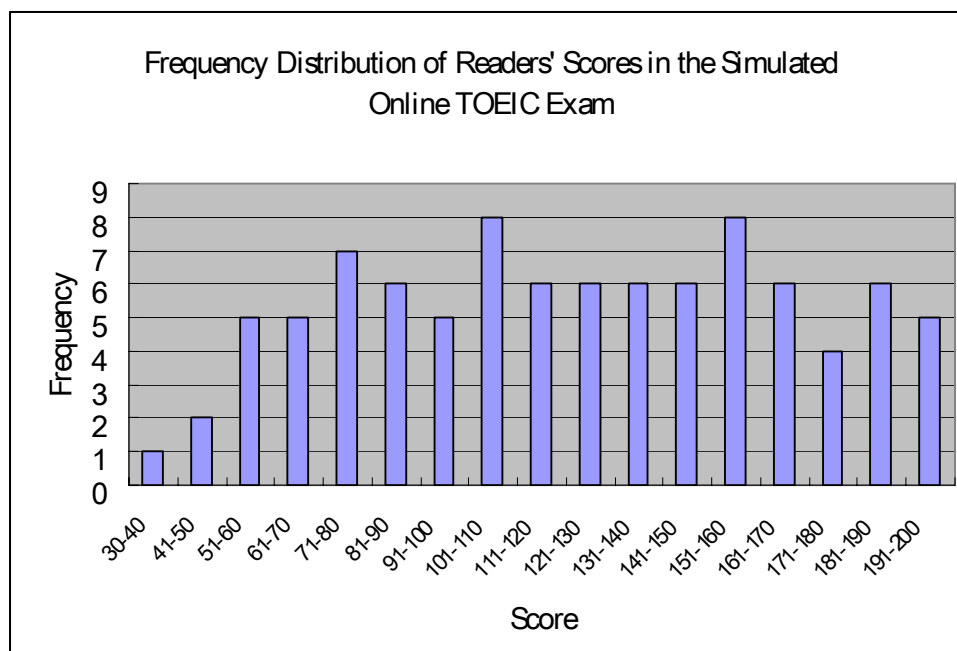


Figure 1. The frequency distribution of the participants' scores

As shown in Figure 1, the highest frequency falls in the two intervals, 101-110 (8 students) and 151-160 (8 students). These two intervals provide the benchmarks for dividing the participants into three groups of readers. Participants with reading scores above 151 were identified as the more-proficient readers and those with reading scores below 101 were the less-proficient readers. Participants with reading scores between 101 and 151 were identified as the average readers.

Thus, 29 more-proficient readers, 32 average readers, and 31 less-proficient readers were identified in this study. The more-proficient readers showed a mean score of 175.52 with a standard deviation of 14.60, the average readers a mean score of 126.41 with a standard deviation of 15.09, and the less-proficient readers a mean score of 74.03 with a standard deviation of 17.53.

## Material

The online referential resolution practice used four texts to examine the participants' reading comprehension. The four texts were selected from *College Reading Workshop* (Malarcher, 2005) based on the following four criteria: abundance of references for reading practice, similar length, similar readability level, texts written for EFL college students. The four texts were *Ideas about Beauty* (number of words: 582; number of referring words: 25), *Fast Food and Teen Workers* (number of words: 583; number of referring words: 25), *Adventure Tours for Charity* (number of words: 599; number of referring words: 43), and *Traditional Markets vs. Modern Markets* (number of words: 577; number of referring words: 31).

## System development

The system built for this study includes three modules, a *user interface*, a *recording module*, and a *feedback module*. Figure 2 shows the interface and module of referential resolution.

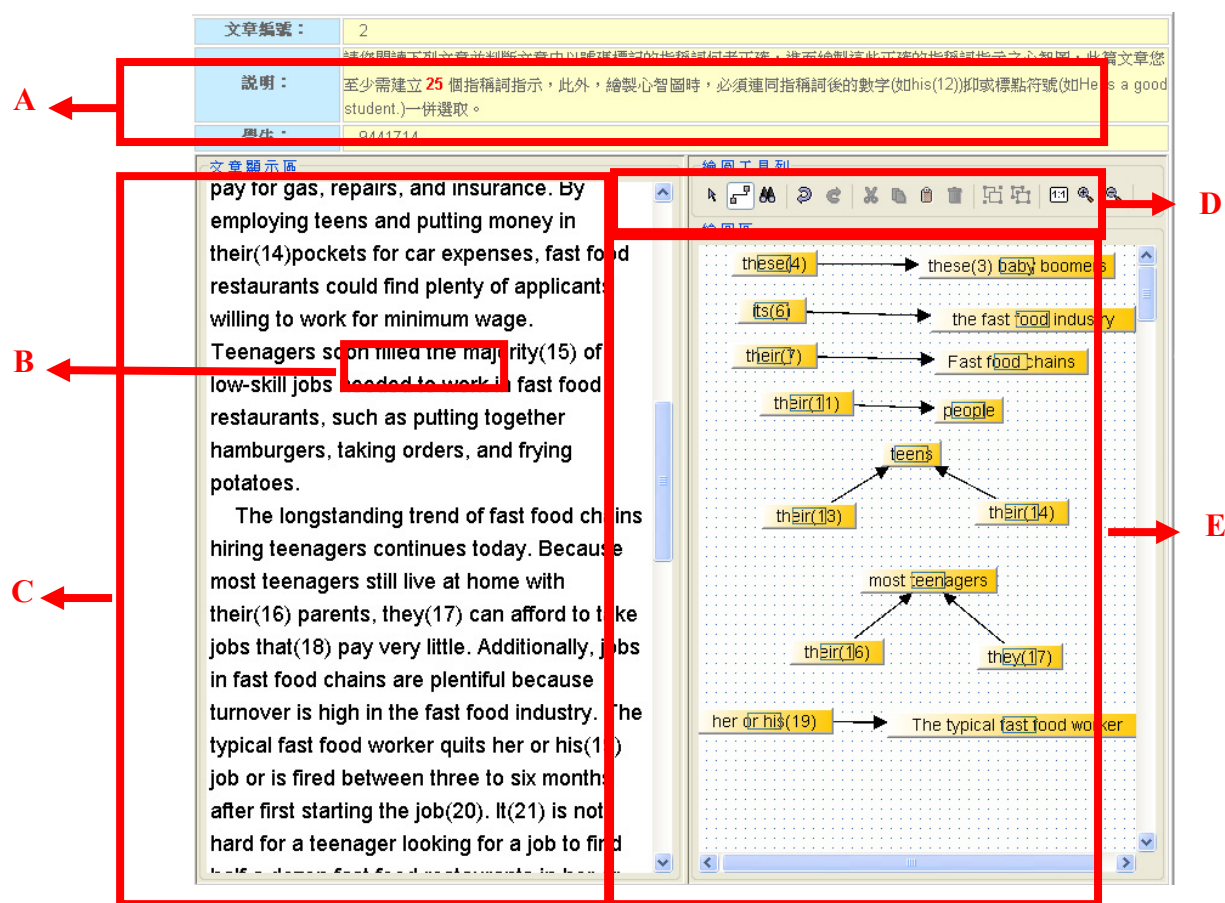


Figure 2. The interface and modules of referential resolution

As shown in Figure 2, the interface and modules of referential resolution include five sections. They are described and shown in details below.

A. *Instruction*. There is a detail description of the referential resolution task. The student has to first identify the references and then resolves them by drawing a mental map to establish the relationship between the references. The instruction prepares and guides the student to approach the task.

B. *References*. All references are marked with numerical numbers. The student has to distinguish the accurate references from the distracters which are not references. He/she then decides what these references refer to, and clicks and drags them to the *canvas* to construct a mental map that illustrates the relationships between the references. When a reference is selected, it is highlighted in the *text field*.

C. *Text field*. This area is used to display the text. Students can select a word or a sentence as a text element and then drag to the canvas directly when they comprehend the relations between the references. When a sentence is selected, it will be highlighted.

D. *Toolbar*. This includes many graphic tools. *Connection tool* is used to establish meaningful relations between referential devices. *Feedback tool* compares students' initial graph and that of the expert. It then informs students what has been done correctly in referential resolution and provides students with three candidate references to correct their errors. Students can decide if they want to activate this tool. Other tools are *cut*, *copy*, *paste*, *erase*, *group*, *ungroup*, *zoom in*, *zoom out*, *undo*, and *redo*.

E. *Canvas*. Students add links to indicate the relationships between references on this canvas. They can *add*, *erase*, *drag* and *drop* elements on the canvas.

## User Interface

The *user interface* includes a *teacher interface* and a *student interface*. The *teacher interface* enables the teacher to manage course data, provide required reading texts, and observe students' reading process and behavior. The *student interface* provides space for the student to draw a graph indicating the relationships between references and answer an online open-ended questionnaire (see Figure 3).

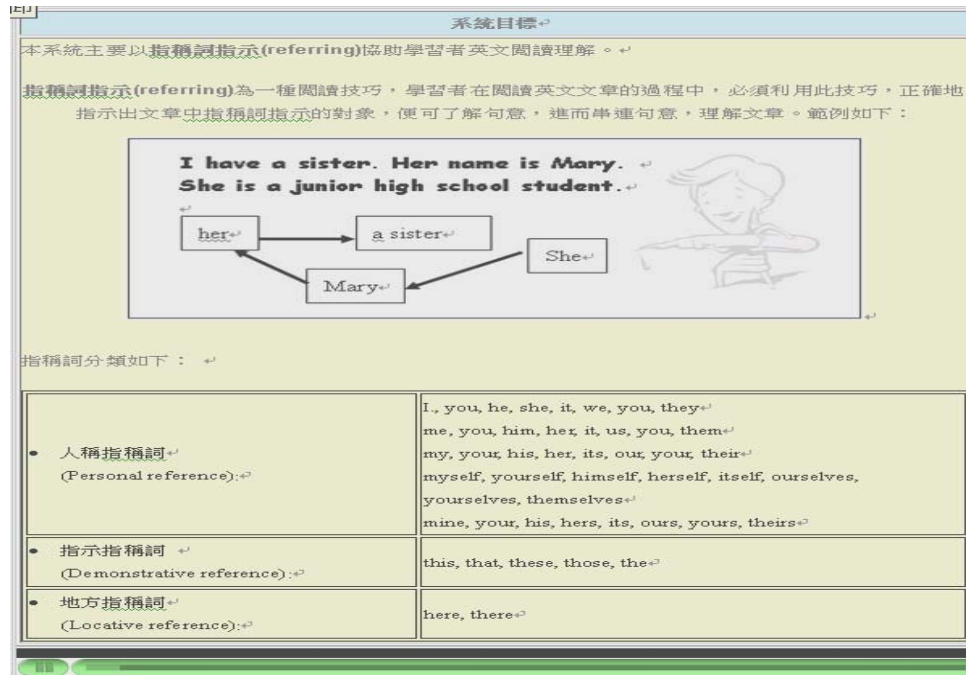


Figure 3. The student interface

## Recording Module

The system uses a *recording module* tracing students' reading process. From the data, the teacher can observe and identify the difficulties students encounter and difference in performance among various reading proficiency groups. The records are helpful for the teacher to modify his/her instruction according to the demonstrated strengths and weaknesses. An example how the *recording module* traces students' reading process is shown in Figure 4.

1:	Read a sentence: [Sigmund Freud ... of the nineteenth century.]
2:	Read a sentence: [He (3) treated ... his (4) "talk cure."]
3:	Select a text element: [Sigmund Freud]
4:	Select a text element: [He (3)]
5:	Select a text element: [his (4)]
6:	Make a connection: [He (3)] connect to [Sigmund Freud]
.....	
38:	Read a sentence: [Freud wrote a book ... dreaming in 1899.]
39:	Select a text element: [his (22)]
40:	Make a connection: [his (22)] connect to [Sigmund Freud]

Figure 4. Records of a student's reading process

## Feedback Module

While students encounter difficulties in constructing their initial maps, they can request feedback. The *feedback module* compares the initial map with the correct one. The results will then indicate those references which are incorrect and offer three candidate references for each incorrect answer. Figure 5 shows the feedback received by the student.



文章編號：	1
說明：	請您閱讀下列文章並判斷文章中以號碼標記的指稱詞何者正確，進而繪製這些正確的指稱詞指示之心智圖，此篇文章您至少需建立 25 個指稱詞指示，此外，繪製心智圖時，必須連同指稱詞後的數字(如his(12))抑或標點符號(如He is a good student.)一併選取。
學生：	a

文章顯示區

agree on whether the factors(4) which influence how most people judge beauty come from genetics (nature) or culture (society).

Devendra Singh, a psychologist at the University of Texas at Austin, conducted an experiment in 1993 to find out if different men found different female body shapes attractive. Dr. Singh gave drawings of different female body shapes to a variety of men and asked them(5) to choose the most attractive body shape. Even though the men(6) came from a wide range of cultural backgrounds, they(7) all tended to rate the "hourglass" body shape as the most attractive. In fact, Dr. Singh found that(8) any woman whose waist is 70% as wide as her(9) hips is judged as attractive by most men no matter how big the woman(10) is overall. Body shape, not weight

(The text is copied from COLLEGE READING WORKSHOP, 2nd edition, Malarcher, 2005)

送出 返回

繪圖工具列

繪圖區

Figure 5. The feedback to student

文章編號：	1
說明：	請您閱讀下列文章並判斷文章中以號碼標記的指稱詞何者正確，進而繪製這些正確的指稱詞指示之心智圖，此篇文章您至少需建立 25 個指稱詞指示，此外，繪製心智圖時，必須連同指稱詞後的數字(如his(12))抑或標點符號(如He is a good student.)一併選取。
學生：	a

文章顯示區

Devendra Singh, a psychologist at the University of Texas at Austin, conducted an experiment in 1993 to find out if different men found different female body shapes attractive. Dr. Singh gave drawings of different female body shapes to a variety of men and asked them(5) to choose the most attractive body shape. Even though the men(6) came from a wide range of cultural backgrounds, they(7) all tended to rate the "hourglass" body shape as the most attractive. In fact, Dr. Singh found that(8) any woman whose waist is 70% as wide as her(9) hips is judged as attractive by most men no matter how big the woman(10) is overall. Body shape, not weight, seemed to be viewed as the critical factor(11) for attractiveness by men in this(12) survey.

(The text is copied from COLLEGE READING WORKSHOP, 2nd edition, Malarcher, 2005)

送出 返回

繪圖工具列

繪圖區

them(5) → a variety of men



Figure 6. The mental map the student draws after feedback



As shown in Figure 5, the three candidate references are highlighted for the student to make a second attempt at the correct answer. Then, the student connects the chosen reference to the referent in the mental map (see Figure 6).

The comparison of the student's answer and the expert's referential resolution result is processed in the following way. First, the module transforms the correct graph and the student's graph into predicates. Next, the module lists all references from the student's predicates and compares them with the correct ones. When the student's references do not match the correct ones, the module will provide one correct reference and two distracters as candidate references for a new attempt at the right answer. Table 1 shows the comparison in the system.

Table 1. The comparison between the student's graph and the correct framework

Correct Map		Student's Map	
Map	Predicates	Map	Predicates
	Add a cell: [A] Add a cell: [B] Referring: [A][B]		Add a cell: [A] Add a cell: [C] Referring: [A][C]

In the table, the *feedback module* compares a student's graph with that of the expert and identifies the resolution of the referential device A to be incorrect. The module will then offer the correct reference B and two distracters D and E as candidate references for the student to make a new attempt.

### Procedures of Data Collection

The online system of referential resolution required students to follow these steps in reading a text: (1) to identify references, (2) to draw relationships between the references, and (3) to answer an open-ended questionnaire. The present study, conducted between October 2<sup>nd</sup>, 2006 and January 3<sup>th</sup>, 2007, involved 92 college students who completed the online referential resolution practice on the website <http://140.125.32.127/reading> in class. The system recorded the reading behavior and process of participants.

### Procedures of Data Analysis

Content analysis and constant comparative method were used to examine the data on participants' mental maps of referential resolution, the *feedback tool* request, and the trace results of their reading process. Content analysis helped the researchers to discover and describe the focus of individual, group, institutional, or social attention and allowed the researcher to make inferences (Weber, 1990; Patton, 1990). Four steps of content analysis were conducted in this study; these were coding, categorization, description, and interpretation. The researchers first read through the transcription and identify the meaning units that can be sentences or passages. Next, the researchers tried to categorize the meaning units and build a system of coding categories. According to categorization, the researchers structure and present the data. The researchers finally interpreted the data by offering explanation, drawing conclusions, and making inferences.

In addition to content analysis, constant comparative method was also used to analyze the data. Comparing incidents applicable to each category was the first stage in which the researchers coded the data into categories and constantly refined the idea of the category as more data were recruited. The second stage was integrating categories and their properties. That is, the researchers focused on the properties of each category and continuously compared the properties among the categories. Then, the comparison enabled the researchers to have theoretical senses and generate the resulting theory. The third stage was delimiting the theory in which the researchers modified and concluded with a theory after comparing the categories. Then, the researchers revised the theory by stripping irrelevant properties of category away and integrating details of properties into an outline of interrelated category. A number of categories was reduced. The final stage was writing the theory in which the researchers finished categorizing and coding data, and developed a theoretical model. Analysis of data using the two research methods will be further explained in the following sections.

## Result

The mental maps were drawn to illustrate the readers' referential process and text comprehension. The analysis of mental maps showed the mental process of different readers with regard to the three types of reference investigated in this study: personal, demonstrative, and locative references. The differences in mental maps of the three types of references are described as follows.

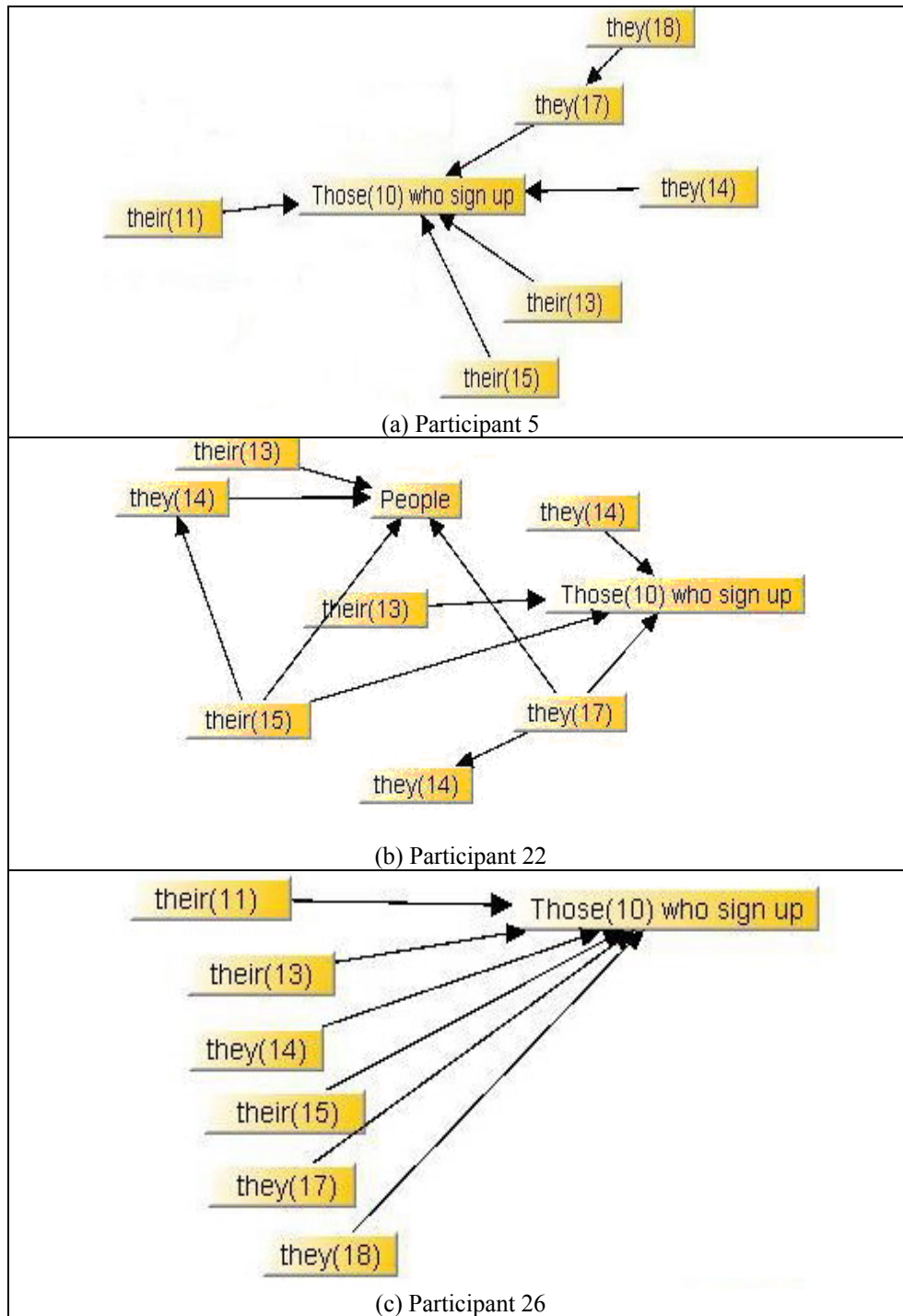


Figure 7. Examples of more-proficient readers' mental maps in personal reference

## Mental Maps of Personal Reference

In the system, participants were asked to read texts and draw their mental maps. A text is shown below.

*Those(10) who sign up then have to pay a deposit in order to hold their(11) space on the tour(12). People cannot get their(13) deposit back later if they(14) change their(15) mind and decide not to go on the tour(16), so they(17) had better be sure they(18) really want to go. The deposit(19) is usually between £200-300.*

Based on the above text and the mental maps that students drew, it was found that the mental process of more-proficient readers in resolving the personal references constructed a network. The network shows a complete and coherent mental representation of textual information. This is illustrated in Figure 7.

In graph (a) and graph (c) of Figure 7, participant 5 referred the references directly to the same subject. Participant 22, in graph (b), further established the relationships among the references. She resolved items *their* (13), *they*(14), *their*(15), *they*(17) and *Those*(10) by associating *they*(17) and *their*(15) with *they*(14), *Those*(10) who sign up, and *people* to establish accurate relationships among these references and referents. In these three examples, networking helped the readers grasp the main idea which is necessary to the correct interpretation of subsequent paragraphs. For example, the trace results of participant 22 further show correct interpretation of follow-up references.

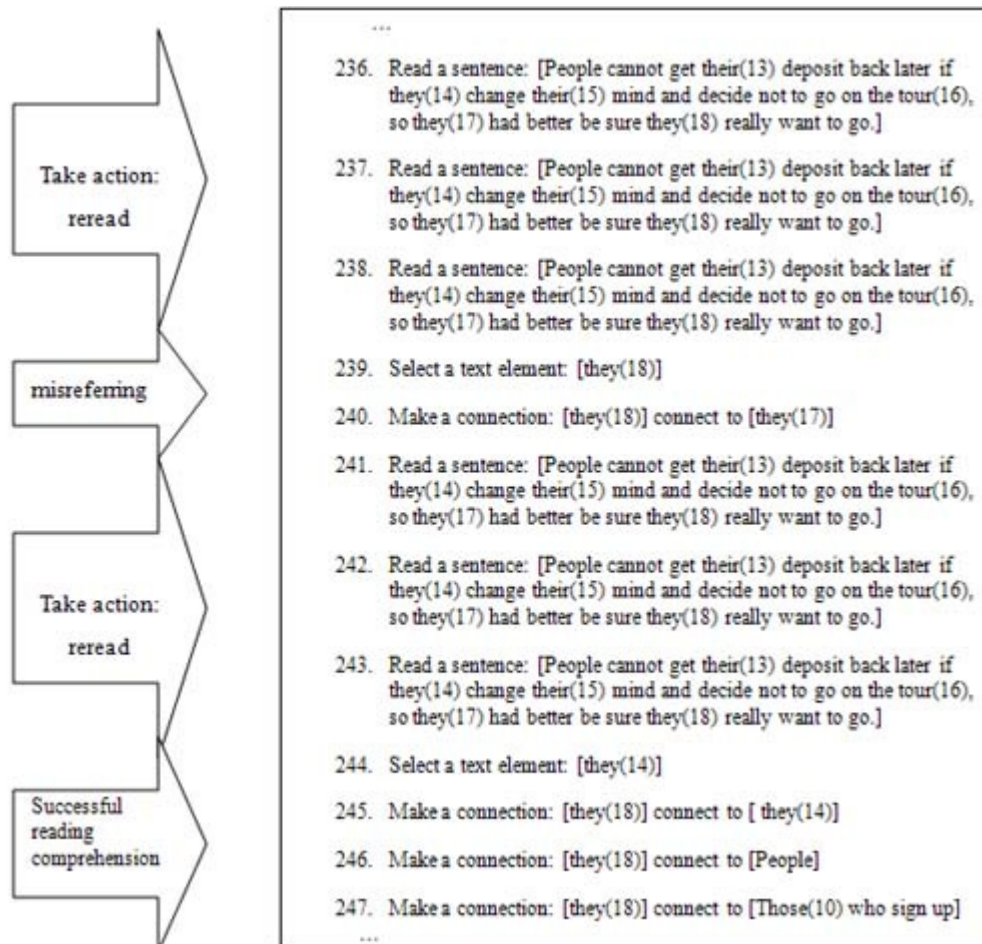


Figure 8. Trace results of participant 22's reading process

As shown in Figure 8, participant 22 read and reread the follow-up sentences for several times (e.g. lines 241~243). She then correctly referred *they* (18) to all the possible subjects in the text (e.g. lines 248, 255, 256, 257) as shown in her mental maps in Figure 9.

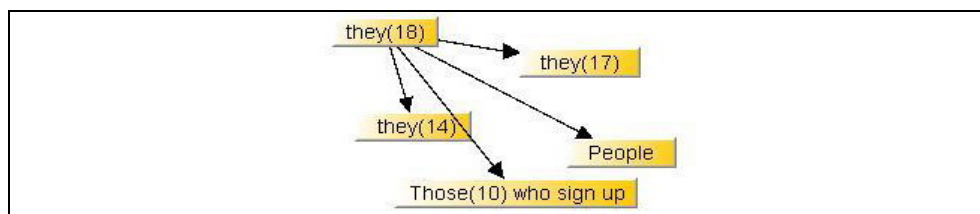


Figure 9. Participant 22's mental map of subsequent reference

In contrast to the coherent network of more-proficient readers, the average and the less-proficient readers failed to develop organized maps. Figure 10 shows the mental maps of average and the less-proficient reader in personal reference.

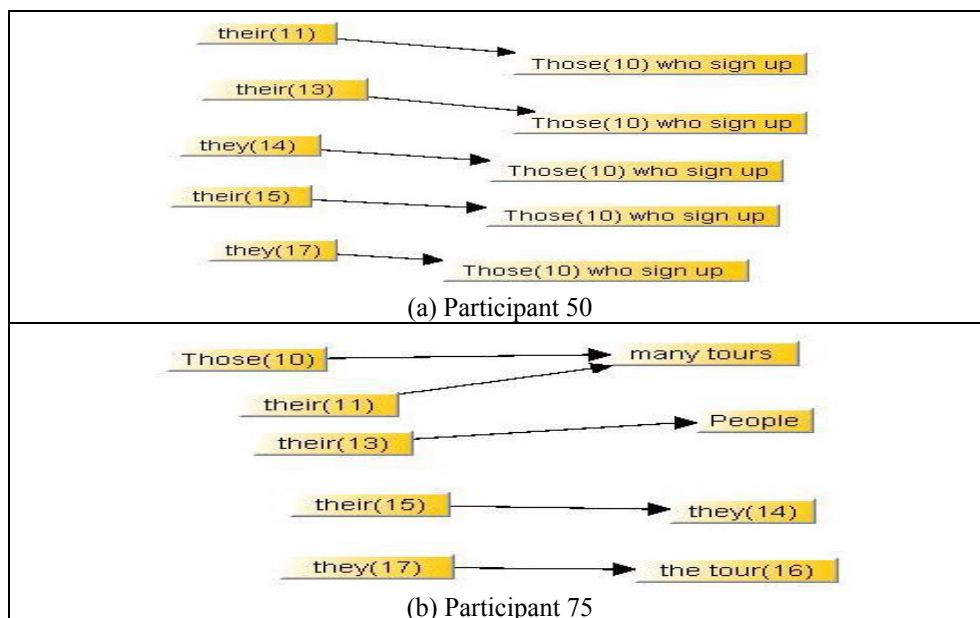


Figure 10. Examples of the average and less-proficient readers' mental maps in personal reference

Graph (a) of Figure 10 indicates that the average reader failed to relate the personal reference to the same subject. She often resolved the references individually, which might explain for her partial understanding of textual information. This also indicated that her textual comprehension was limited to a single sentence or paragraph.

As for the less-proficient reader, he not only referred the references to a subject separately but also mismatched the references with an incorrect subject. Graph (b) of Figure 10 illustrates his disassociation of plural words to the single subject. He mismatched *they* (17) with the former word, *the tour*(16). The separation or mismatch in referential resolution led the less-proficient reader to misinterpretation of the current sentence and the subsequent text. This resulted in overall miscomprehension of the text. For instance, Figure 11 shows participant 75's mismatch of the subsequent reference.

As shown in Figure 11, participant 75 requested the *feedback tool* twice for referring *this*(22) to a subject (e.g. lines 74 and 76). However, he still mismatched *this*(22) with *1,500-2,500*. Similar comprehension failures occurred repeatedly (e.g. lines 86~90).

### Mental Maps of Demonstrative Reference

A text was also read by three groups of participants and it is shown below.

*Through his(21) own survey, Dr. Yu found that(22) the men(23) in this(24) isolated community preferred heavier women with wider waists, and not particularly women with "hourglass" shapes.*

Because this(25) small community has lived apart from western mass communication, their own culture has not been influenced by outside standards of beauty. Dr. Yu points out that(26) this(27) group has experienced the same genetic evolution as all humans do, but a different standard. In order to check the reliability of his(28) study, Dr. Yu surveyed two other groups of men from this(29) same community. However, the second and third groups surveyed by Dr. Yu had had more exposure to Western media. The results of these(30) later surveys showed that(31) as men from this(32) isolated community came into contact with Western media, their(33) standards of beauty began to change more toward the Western standard of beauty.

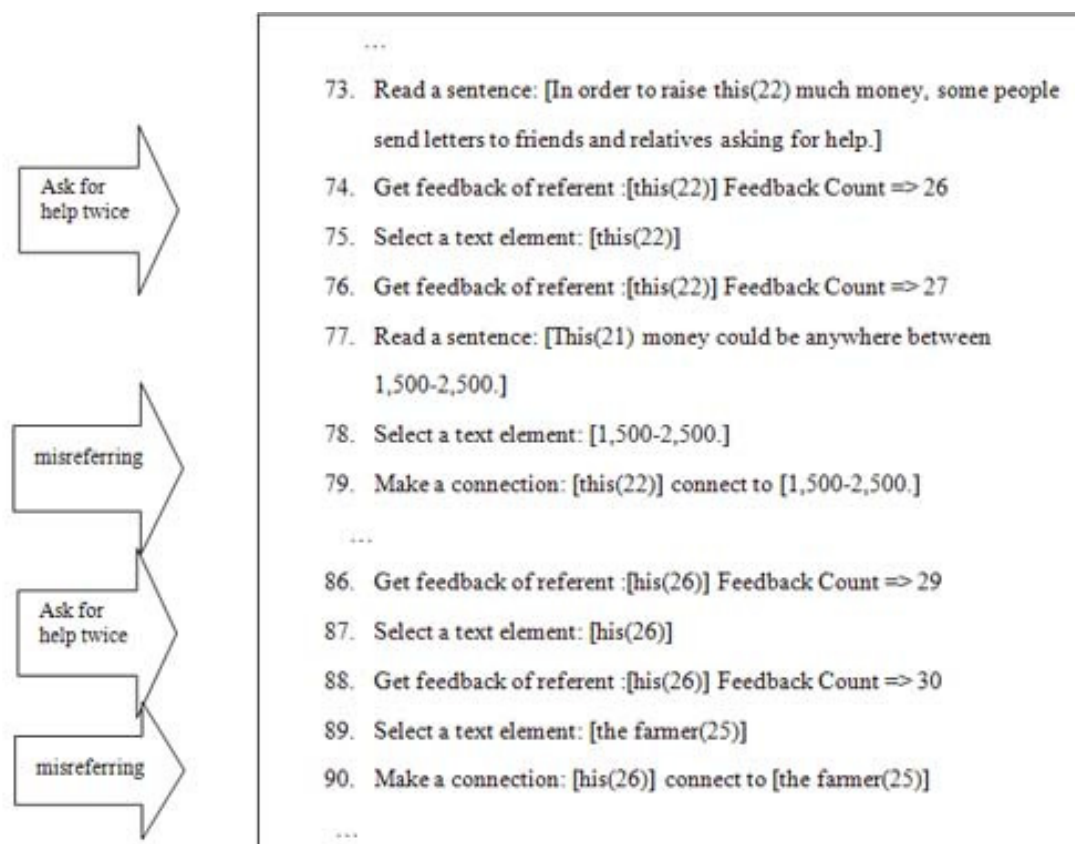


Figure 11. The trace results of participant 75's reading process in the subsequent sentences

While reading the above text, readers were drawing their mental maps. From the analysis of the readers' mental maps, it was found that the more-proficient readers successfully related the demonstrative references in different parts of the text to the same subject in an integrative way. Examples of the more-proficient readers' mental maps are presented in Figure 12.

As shown in graph (a) of Figure 12, participant 23 successively referred the demonstrative references to the same subject. In graph (b), participant 26 even correctly established the relationships between references. She correctly linked *the shopper*(31) to the former reference, *the shopper*(29). A clear understanding of the relationships among references and textual information aided the more-proficient readers to resolve lexical ambiguities and better comprehended the text. This supported her to refer the subsequent references to the correct subjects. Figure 13 presents participant 26's mental map of follow-up sentences.

In Figure 13, participant 26 correctly referred and integrated the references. Her successful resolution of references in previous sentences facilitated her correct interpretation of follow-up references and constructed relationships among sentences.



Conversely, the average readers commonly referred the demonstrative references either to part of the subject or incorrect ones and the less-proficient readers often mismatched the references to the incorrect subject. Examples of average readers' mental maps are shown in Figure 14.

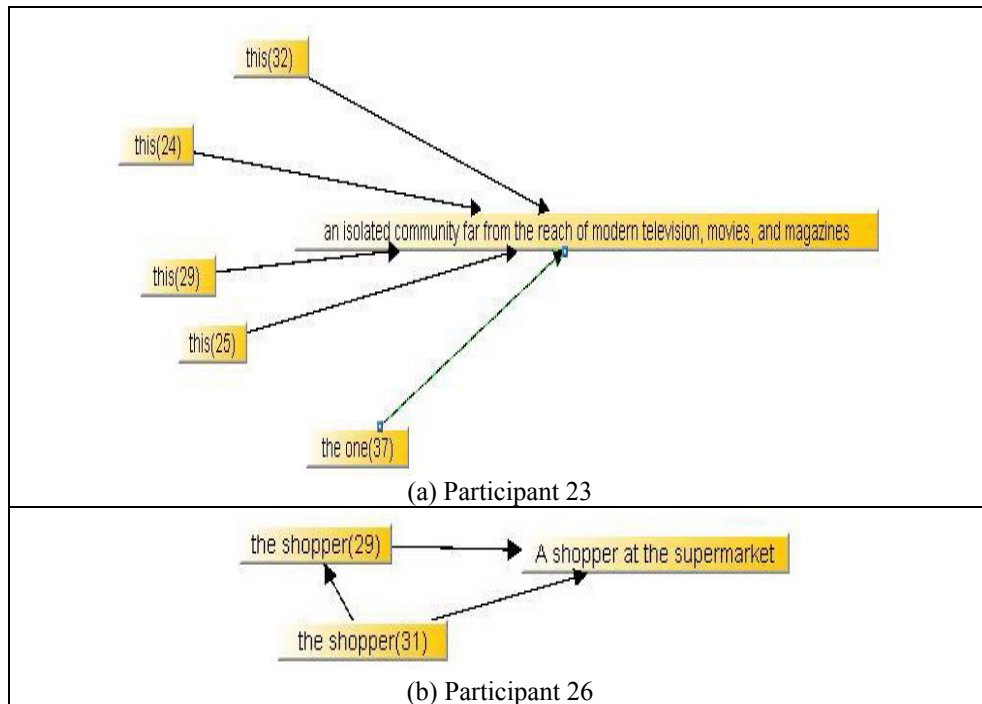


Figure 12. Examples of the more-proficient readers' mental maps in demonstrative reference

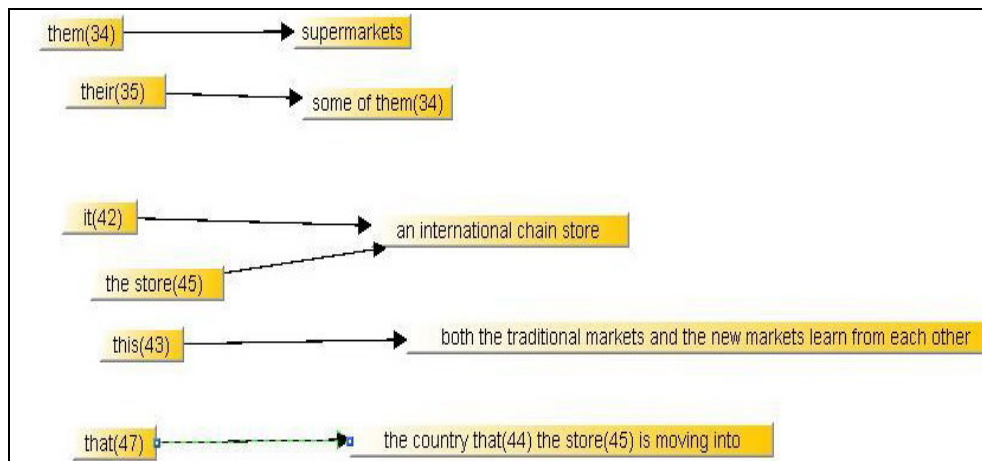


Figure 13. Participant 26's mental map of follow-up references

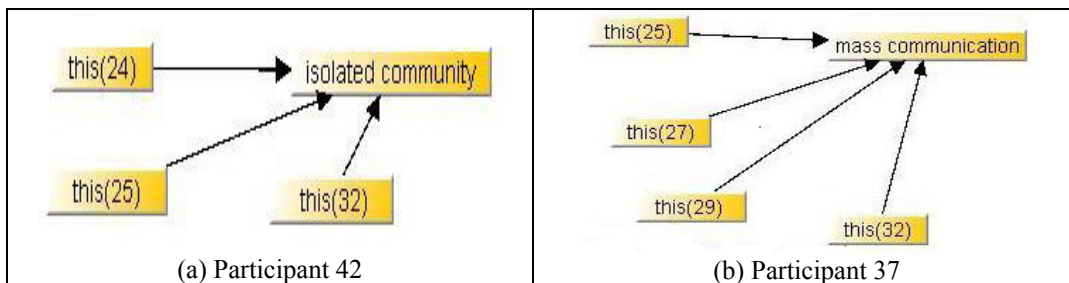


Figure 14. Examples of average readers' mental maps in demonstrative reference

Graph (a) of Figure 14 indicates that the average reader could not match the references to a correct subject. Instead, she could only refer the references, *this* (24), *this*(25), and *this*(32) to the partial subject, *isolated community*. In graph (b), another average reader even linked the references to the incorrect subject. He mismatched the demonstrative references to *mass communication*. The incorrect or partial integration of references resulted in misunderstanding of these parts of the text, which, in turn, hindered comprehension of related information in the succeeding sentences or paragraphs. Figure 15 shows participant 42's mental map of follow-up references.



Figure 15. Participant 42's mental maps of follow-up references

In Figure 15, participant 42 constantly mismatched the references with incorrect subjects. *These*(34), which is a demonstrative reference referring to objects or events in the context, was matched with *men from this(32) isolated community came into contact with Western media, their(33) standards of beauty began to change more toward the Western standard of beauty*, which is a personal pronoun indicating an individual. *It*, which is a personal pronoun, was mismatched with *find isolated communities*, which is a verb phrase.

As for the less-proficient readers, they often referred several occurrences of the same referential word incorrectly to one subject. In graph (a) of Figure 16, the reference *this* in different parts of the texts was incorrectly referred to the subject *southeast Peru*. This incorrect referential resolution led the less-proficient reader to misinterpret *southeast Peru* and activated irrelevant background knowledge to interpret the subsequent sentence. In graph (b), another less-proficient reader also related the references to incorrect subjects. He incorrectly referred the demonstrative reference *this* (24) to a person referent *heavier women*.

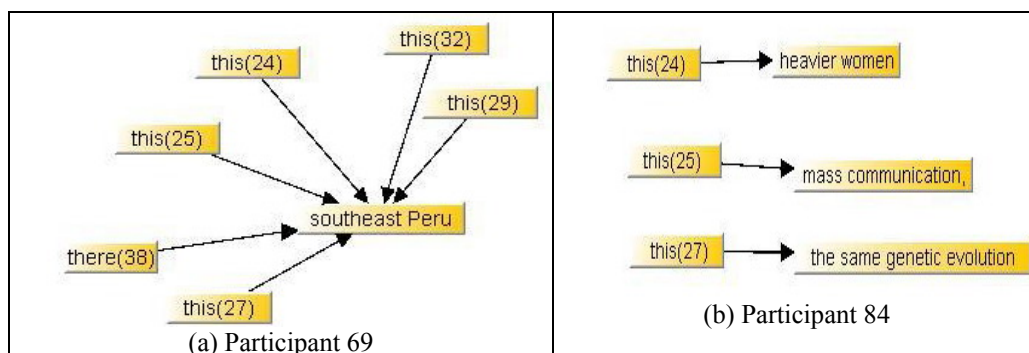


Figure 16. Examples of the less-proficient readers' mental maps in demonstrative reference

The mismatches shown in Figure 16 confused the reader in the organization of textual information, either by leading comprehension in the incorrect way or by misinterpreting the meaning of textual information. Figure 17 presents participant 69's mental map of follow-up references.

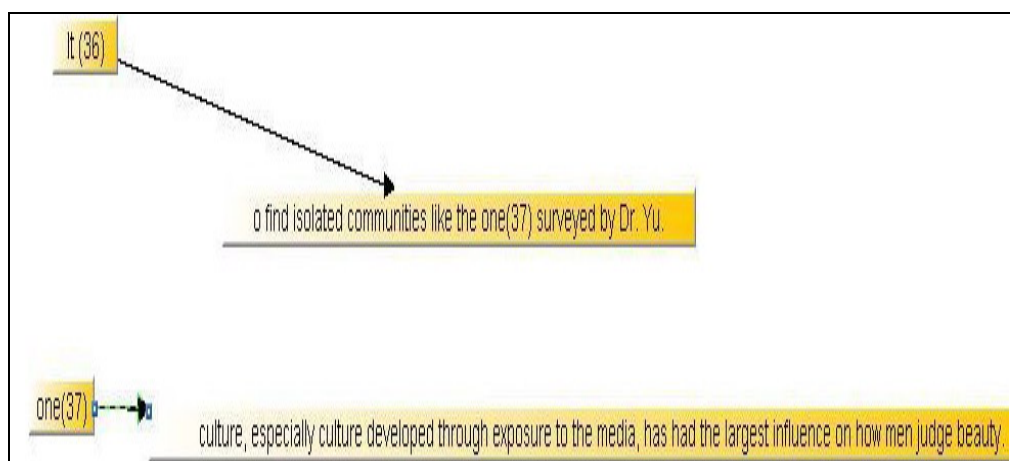


Figure 17. Participant 69's mental maps of follow-up references

In Figure 17, participant 69 not only failed to identify the correct references but also mismatched references with incorrect subjects. *One (37)*, which was not a reference, was selected and mismatched. *It (36)*, which is a personal pronoun, was connected to *o find isolated communities like the one (37) surveyed by Dr. Yu.*, which is a verb phrase.

### Mental Maps of Referring in Locative Reference

From mental maps of locative reference, the more-proficient readers were able to identify the references and referred them to the correct subject (see graph (a) of Figure 18) in reading the following text.

*One reason ferias are still popular is because people feel that(19) they(20) can get better fruits and vegetables at a better price there(21). While shopping at the feira, people can pick up the things they(22) might want to buy, smell the fruits and vegetables, and sometimes even taste them(23) before they(24) buy anything.*

The correct identification and resolution facilitated the more-proficient reader to accurately interpret follow-up textual information and fully comprehend the text. The average readers, however, tended to refer the references to the incorrect subject due to their inaccurate comprehension of the text. In graph (b) of Figure 18, the average reader (participant 50) mismatched *there (21)* with *the feira* when *There (21)* is not associated with a specific feira. Hence, the average reader referred the locative reference to an incorrect semantic subject.



Figure 18. Examples of the more-proficient and average readers' mental maps in locative reference

Similarly, the less-proficient readers were deficient in resolving locative reference. For example, in graph (a) of Figure 19, the less-proficient reader referred *ferias* to *there (21)*. It showed that the reader first recognized the word “*ferias*” and then linked the word to the reference *there (21)*. This mental process suggested that the reader had to spend time and effort on word recognition first. If she encountered unknown words in the following sentences, she would fail to refer the references correctly.

In addition, graph (b) of Figure 19 indicates that the less-proficient reader could not correctly identify the locative reference. She was confused by the references *the country (17)* and *there (21)* and referred them to the same subject. This misinterpretation continued in follow-up sentences that contained the two references. In graph (c), the less-proficient reader referred the locative reference to an incorrect subject. *There (21)*, which refers to a place, was mismatched with *a better price*, a noun indicating an amount of money.



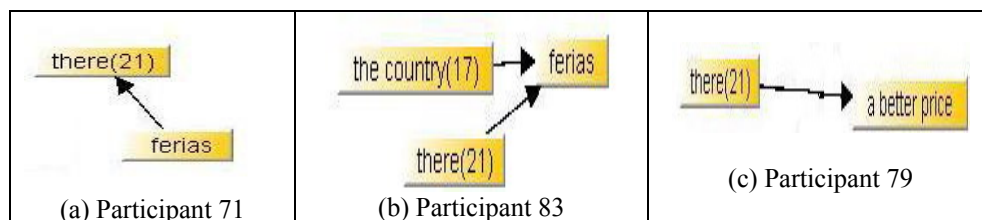


Figure 19. Examples of less-proficient readers' mental maps in locative reference

The mismatch of earlier references could lead to comprehension failures of the subsequent text. For instance, participant 83 continuously matched the references in follow-up sentences to incorrect subjects, linking *this* (28) and *their* (35) to incorrect subjects (see Figure 20).

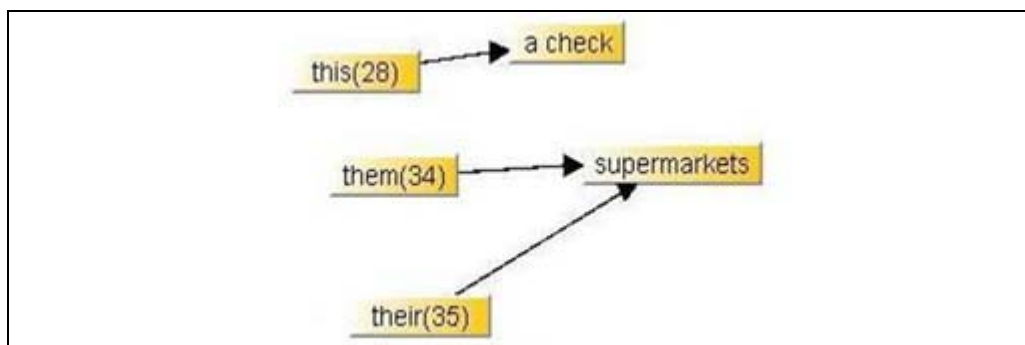


Figure 20. Participant 83's mental maps of follow-up references

### An open-end questionnaire

The open-ended questionnaire was designed to understand readers' perception toward referential resolution and the online system. Before using the system, many readers stated that they did not know the referring strategy and seldom used it. The more-proficient readers expressed that they were only familiar with the pronouns. The less-proficient readers claimed that they seldom identified references in a text. After using the system, most readers stated that the system enhanced their knowledge and employment of referring strategy. In the following tables, MR refers to the more-proficient reader, AR the average reader, and LR the less-proficient reader. Text 1 indicates the text *Ideas about Beauty*, Text 2 *Fast Food and Teen Workers*, Text 3 *Adventures Tours for Charities*, and Text 4 *Traditional Markets and Modern Markets*.

When participants were asked their reading difficulties and solutions in the online system, some more-proficient readers expressed that they sometimes failed to identify references in a text (see participant 21's response in Table 2). This was the same for the average and less-proficient readers who were unable to resolve the references in longer sentences (see participant 44's and 62's responses in Table 2). To overcome their difficulties, they read and reread related sentences to find out the contextual clues. They even requested the *feedback tool* for assistance. For instance, when reading the first text, participant 44 read and reread sentences for contextual clues. After this, she began to reflect on the possible solutions in the second and third texts. In the fourth text, she monitored her reading process and requested the *feedback tool* very often for assistance.

Participants commonly emphasized the importance of the feedback tool for their reading progress. For instance, the more-proficient readers relied on the feedback tool to verify whether the connections between references were correct in their mental maps (see participant 25's response in Table 3). Both the average and the less-proficient readers asked help from the feedback tool to identify references and figure out the relationship between two words (see participant 41's and 62's responses in Table 3). For example, participant 25 began to clarify her referential resolution by connecting the references in the second text. After this, she sought help from the feedback tool to confirm the connections she made in the third and fourth text.

Table 2. Participants' reading difficulties and resolutions in the system

Group	Text	Response
MR	1 P21	If I couldn't identify the reference, I reread related sentences or making inference based on the contextual clues. In addition, I requested the <i>feedback tool</i> for help.
	2 P21	I reread the sentences to better understand their meanings so that I could identify references in a text.
	3 P21	I did not have difficulty in using the online system.
	4 P21	When I encountered difficulty, I requested the <i>feedback tool</i> for assistance.
AR	1 P44	I could do well most of the time. However, I sometimes failed to identify references. I resolved these difficult references by rereading the text and by contextual clues.
	2 P44	Sometimes I resolved the references incorrectly, such as selecting incomplete statements of a subject or adding too many textual elements as a subject. After I was informed of the errors, I would reread the text.
	3 P44	I have no difficulty.
	4 P44	I would try to resolve the references first and then requested the <i>feedback tool</i> .
LR	1 P62	I was confused by the long sentences in a text. If I was unfamiliar with the online system, I would consult my peers.
	2 P62	Since I had solved the technical problems that I encountered in the first text, I did not encounter the same problems in the second text. By this online practice, I realized that I am poor at referential identification and resolution. I needed to spend longer time on solving references.
	3 P62	I do not have difficulty using the system. The only difficulty I encountered was that there were too many references in a text. I think that I did not have solid knowledge and strategy usage in resolving references.
	4 P62	The only difficulty that I could think of is that I need to spend longer time on figuring out the relationships between words.

Similar to the more-proficient reader, the average reader also expressed her perception in using the feedback tool. She indicated that the feedback tool assisted her to narrow down the possible answers when reading the first text. With the assistance of the feedback tool, she began to identify more references in the second text. In the third and fourth text, she was able to find out the correct references (see participant 41's responses in Table 3).

The less-proficient reader's perception toward the feedback tool is also illustrated in participant 62's responses in Table 3. When reading the first text, she relied on the feedback tool to identify and refer the references. In the second and third texts, she was able to identify and refer the references based on the clues that the feedback tool offered. In the fourth text, she emphasized that the feedback tool aided her to identify and refer the references in longer texts.

Table 3. Participants' perception towards the *feedback tool*

Group	Text	Response
MR	1 P25	If I was confused by referential resolution, I requested the <i>feedback tool</i> . Then, I knew what the reference refers to.
	2 P25	The feedback tool helped me solve my problems .
	3 P25	It enabled me to confirm my reference identification.
	4 P25	It helped me detect whether I made an error or not.
AR	1 P41	It helped me narrow down the possible answers.
	2 P41	It helped me identify the correct references.
	3 P41	It aided me to find out the references.
	4 P41	It helped me find out more references in a text.
LR	1 P62	It enabled me to rebuild my confidence in the referring tasks.
	2 P62	It is helpful.
	3 P62	It provides me with candidate words.
	4 P62	It is very helpful, particularly for longer texts.

When participants were asked how the referential resolution influenced their reading comprehension, most readers mentioned that it helped them reread a text and grasp the main idea. For example, participant 21 in Table 4 mentioned that referential resolution assisted her to grasp the main ideas and figure out the meanings of unfamiliar words in the first text. In the second and third text, she was able to read the texts faster. In the fourth text, she better comprehended the text and enlarged her vocabulary size. The average reader, participant 33's response in Table 4, expressed that referential resolution was not very helpful in reading the first text. In the second text, he began to realize that it assisted him to understand the textual information and read faster. In the third text, he was able to grasp the main ideas. In the fourth text, he could correctly interpret the text. The less-proficient reader, participant 68's response in Table 4, was also benefited from referential resolution. In the first and second texts, he claimed that it helped him understand the textual information. In the third text, he began to figure out the unfamiliar words by resolving the references. In the fourth text, he was able to understand the theme of the text (see participant 68's response in Table 4).

*Table 4. Students' perception towards the influence of referential resolution on reading comprehension*

Group	Text	Response
MR	1 P21	It helped me understand the full text by grasping the main ideas and enlarging vocabulary size.
	2 P21	It improved my reading comprehension and speeded up my reading pace
	3 P21	It aided me to read comprehensively.
	4 P21	I could have deeper understanding about the text. I also enlarged my vocabulary size and better my reading pace and comprehension.
AR	1 P33	I thought it was not very helpful.
	2 P33	It improved my reading pace and better comprehended the text.
	3 P33	I could quickly grasp the main ideas of a text.
	4 P33	I could fully comprehend a text.
LR	1 P68	It helped me better understand the textual information.
	2 P68	It aided me to understand the text.
	3 P68	It helped me to figure out the meanings of the unfamiliar words.
	4 P68	It helped me understand the theme of the text.

The readers' responses in the open-ended questionnaire indicated the different perceptions among the readers. It was found that the referential resolution facilitated the more-proficient readers to read faster and fully comprehend the text whereas it aided the average readers to grasp the main ideas. The less-proficient readers were particularly assisted to understand the meanings of unfamiliar words since they encountered greater difficulties in vocabulary than the more and average readers did. The readers' different perceptions revealed their various difficulties in reading texts.

## Conclusion

From the analysis of college students' mental maps, reading process data, and perception in an open-ended questionnaire, several facts were found. First, the online system guided students to be engaged in comprehension monitoring and referring strategy when reading texts. To draw mental maps, the more-proficient students read and reread the previous sentence and subsequent ones repeatedly. In rereading, they were actually engaged in comprehension monitoring to establish the relationships between the former reference and the follow-up one. The understanding of relationships among the references and the textual information further aided them to resolve lexical ambiguities and grasp main idea. The average and less-proficient readers, in contrast, very often referred the references to partial subjects or incorrect ones. The mismatch of references caused their partial understanding of the text or complicated their text interpretation. This hindered their text comprehension as well as the storage of information in memory. It should be emphasized to these readers the benefits of using feedback tools and referring strategy.

Second, readers' responses in the open-ended questionnaire also indicated their progress in referential resolution. For example, in the first text, the less-proficient reader had difficulties reading long sentences and using the computer system. She asked help from her peers to overcome her difficulty in using the system first. In the second text, she

only had problems in identifying correct references. In the third text, she developed her referential identification and resolution in reading but failed to this task when there were too many references in a text. In the fourth text, she continued to develop her referential identification and her difficulty was changed to referential resolution. She thought that she had to spend longer time on figuring out the relationship between two words.

Third, the computer system aided the teacher to identify students' reading difficulties. In referential identification, the teacher was able to observe which reference items and types were difficult for his/her students. In referential resolution, for instance, the teacher could monitor students' process and progress in establishing relationships between references through recorded data in the *trace result*. This recorded every action that the student took during the reading process and made the intangible process visible. It also enabled the teacher to design or modify her/his follow-up instruction to help the students overcome their difficulties and better develop their integrative skills in reading.

Finally, the implication is that reading process shares equal importance as the reading product in the manifestation of readers' abilities. Traditionally, students' performances are evaluated by the reading product as expressed in proficiency levels or comprehension scores. A single level or score does not reveal detail information on reading strengths, weaknesses, and difficulties. It gives the teacher very little clue on the substantial needs of less-proficient readers to design instructional intervention. This perpetuates less-proficient readers' recurrence of comprehension failures in a text or in all reading tasks. As shown in this study, mental maps and reading processes clearly explain students' reading success and failures. The discussion of these maps and processes between the teacher and students will be one of the ways to promote latter's awareness of the need to employ comprehension monitoring and referring strategy.

## References

- Chang, K. E., Sung, Y.T., & Chen, I.D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education*, 71(1), 5-23.
- Chen, C. M., You, T. Y., Yang, Y. F., & Huang, C. C. (2004). *An evaluation of English proficiency tests for college students in Taiwan*. Taiwan: Ministry of Education.
- Dreyer, C. (1998). Improving students' reading comprehension by means of strategy instruction. *Journal for Language Teaching*, 31(1), 18-29.
- Dreyer, C. & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System*, 31, 349-365.
- Greene, J.A., & Azevedo, R. (2007). Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research*, 36(2), 125-148.
- Hartley, K. (2001). Learning strategies and hypermedia instruction. *Journal of Educational Multimedia and Hypermedia*, 10(3), 285-305.
- Hartman, H. J. (2001). Developing students' metacognitive knowledge and skills. In H. J. Hartman (Ed.), *Metacognition in Learning and Instruction: Theory, Research and Practice*, Boston: Kluwer Academic, 33-68.
- Hazzan, O. (2004). Mental Constructions and construction of web sites: learner and teacher points of view. *British Educational Communications and Technology*, 35 (3), 323-344
- Huang, S. H. (2005). *Assessing the Relationship between Referential Understanding and Academic Reading Comprehension among EFL College Students*, Masters thesis. National Yulin University of Science and Technology, Taiwan.
- Levine, A., Ferenz, O., & Reves, T. (2000). EFL academic reading and modern technology: How can we turn our students into independent critical readers? *TESL-EJ* 4(4), Retrieved September 11, 2008 from <http://www-writing.berkeley.edu/TESL-EJ/ej16/al.html>.
- Kwon, S. Y., & Cifuentes, L. (2007). Using computers to individually-generate vs. collaboratively-generate concept maps. *Educational Technology & Society*, 10(4), 269-280.

- Malarcher, C. (2005). *College Reading Workshop*, 2<sup>nd</sup> Ed., New York: Compass.
- Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones, & L. Idol (Eds.), *Dimensions of Thinking and Cognitive Instruction*, Hillsdale, NJ: Lawrence Erlbaum, 15-51.
- Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods* (2<sup>nd</sup> Ed.), Newbury Park, CA: Sage.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human- Computer Studies*, 58 (3), 327-345.
- Pritchard, R. E., Romeo, G. C., & Muller, S. A. B. (1999). Integrating reading strategies into the accounting curriculum. *College Student Reading*, 33(1), 77-82.
- Ruiz-Primo, Schultz, Li, & Shavelson (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, 38(2), 260-278.
- Schacter, J., Herl, H. E., Chung, G. K.W. K., Dennis, R.A., & O'Neil, H.F. Jr. (1999). Computer-based performance assessments: a solution to the narrow measurement and reporting of problem-solving. *Computer in Human Behavior*, 15 (3-4), 403-418.
- Walsh, C. R., & Johnson-Laird, P. N. (2004). Co-reference and reasoning. *Memory & Cognition*, 32(1), 96-106.
- Weber, R. P. (1990). *Basic Content Analysis* (2<sup>nd</sup> Ed.), Newbury Park, CA: Sage.
- Wood, E., Motz, M., & Willoughby, T. (1998). Examining students' retrospective memories of strategy development. *Journal of Educational Psychology*, 90, 698-704.
- Yang, Y. F. (2003). The development of diagnostic reading inventory, *Proceedings of the 20<sup>th</sup> International Conference on English Teaching and Learning in R. O. C.*, 273-288.
- Yang, Y. F. (2004). The pilot study of *Diagnostic Reading Assessment*, *Proceedings of the 21<sup>st</sup> International Conference on English Teaching and Learning in R. O. C.*, 241-254.
- Yang, Y. F. (2007). The prevention of English reading difficulties in Taiwan. *Studies in English Language and Literature*, 20, 75-90.
- Yang, Y. F., Yeh, H. C. & Wong, W. K. (2008). Constructing mental representation of reference by feedback in a computer system. *Computers in Human Behavior*, 24 (5), 1959-1976.

## **A cross-cultural examination of the intention to use technology between Singaporean and Malaysian pre-service teachers: an application of the Technology Acceptance Model (TAM)**

**Timothy Teo\***

Nanyang Technological University, 1 Nanyang Walk, Singapore 637616 // timothy.teo@nie.edu.sg

**Wong Su Luan**

Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia // wsuluan@gmail.com

**Chai Ching Sing**

Nanyang Technological University, 1 Nanyang Walk, Singapore 637616 // chingsing.chai@nie.edu.sg

\*corresponding author

### **ABSTRACT**

As computers becomes more ubiquitous in our everyday lives, educational settings are being transformed where educators and students are expected to teach and learn, using computers (Lee, 2003). This study, therefore, explored pre-service teachers' self reported future intentions to use computers in Singapore and Malaysia. A survey methodology was employed such that validated items from past relevant research work were adopted. Based on the Technology Acceptance Model (TAM), 495 completed surveys of pre-service teachers were collected from both countries. Structural Equation Modeling (SEM) was implemented on the data captured. This study contributes to the growing multi-cultural studies on TAM by demonstrating that perceived usefulness (PU), perceived ease of use (PEU) and computer attitudes (CA) to be significant determinants of both Singaporean and Malaysian pre-service teachers' behavioral intention (BI). Differences were, however, detected between Singaporean and Malaysian pre-service teachers in terms of PU, PEU and CA but no differences were detected in BI with regards to technology acceptance.

### **Keywords**

Technology acceptance, pre-service teachers, cross-cultural, structural equation modeling

### **Introduction**

Researchers have been investigating the factors that drive individuals to adopt and use information technologies in their workplace and personal lives. Consequently, technology adoption and use, often referred to as user acceptance, has become one of the most researched areas in the information science literature (Agarwal, 2000; Hu, Chau, Liu-Sheng & Yan-Tam, 1999, Smarkola, 2007). Research on technology acceptance often include (1) determining the factors that cause individuals to accept or reject new information technology, or engage in sabotage or active resistance, (2) designing appropriate implementation tactics and interventions that mitigate problems associated with the rejection of information technologies, and (3) factors that ensure continual use of information technologies.

### **Use of technology by teachers**

Despite the widespread use of technology in the schools, it was found that teachers' use of computer technology is mostly for administrative support rather than instructive purposes (Becker, 2001). He found that teachers generally used computer technology to support their existing practices (providing practice drills, demonstration) and communication (such as the use of email) rather than to engage students in learning that involve higher order thinking (see also Wozney, Venkatesh, Abrami, 2006). Becker (2001) surveyed more than 4,000 grade 4-12 teachers in over 1,100 schools across the United States of America. Among the major survey findings were teachers' infrequent use of computers in the classroom and occasional assignment to lower-ability classes with computer games and drills. These findings reflected the supportive use of computer technology rather than using it as an

instructional tool for teaching. There are other instances of technology that was not used adequately used in the classrooms. Burns (2002) found that teachers restricted the use of computers in the classroom due to their fears that students might break the computers. These teachers could be resisting change and could not use technology to its fullest potential.

Educational institutions are witnessing a profound increase in the use and range of technology for example, multimedia presentations, video teleconferencing and, more currently, Web-based Instruction. Technology applications in teaching and learning were considered as the wave of the future; they have a direct impact on current educational practices and policies and subsequently have the potential to alter traditional definitions of education (Bereiter & Scardamalia, 2006). It is important for teachers to understand the precise role of technology so that they can effectively cope with the pressure created by continual innovation in educational technology and tensions to prioritize the use of technology. Given the current state of technology use in education, while it is reasonable to expect teachers to believe the impact of technology on students' learning, teachers may or may not accept technology due to personal factors, such as computer efficacy (Gong, Xu, & Yu, 2004), technical factors such as (Thong, Hong, & Tam, 2002) and environmental factor such as facilitating conditions (Ngai, Poon, & Chan, 2007). Therefore, the need to understand technology acceptance by teachers calls for an examination into the factors that influence teachers' acceptance of technology.

## **Literature Review**

### **Technology Acceptance Model (TAM)**

In the last three decades, researchers have given much attention to identifying the conditions or factors that could facilitate technology integration into businesses (Legris, Ingham, & Colletette, 2003). Arising from this motivation, models were developed and tested to help in predicting technology acceptance. Among these models, the Technology Acceptance Model (Davis, Bagozzi, & Warshaw, 1989) is arguably the most popular in technology acceptance studies (McCoy, Galletta & King, 2007). Overall, the TAM has been empirically proven successful in predicting about 40% of a system use (Legris et al. 2003). It has also found to be a parsimonious representation of how perceptions and attitudes affect technology use (Sivo & Pan, 2005).

The origins of the TAM came from Ajzen and Fishbein's (1980) Theory of Reasoned Action (TRA). It posits that beliefs and attitudes are related to individuals' intentions to perform. According to TRA, attitude toward a behavior is determined by behavioral beliefs about the consequences of the behavior (based on the information available or presented to the individual) and the affective evaluation of those consequences on the part of the individual. Beliefs are defined as the individual's estimated probability that performing a given behavior will result in a given consequence.

Introduced and developed by Davis (1989), the TAM is a model that addresses the issue of how users come to accept and use a technology. There are two specific variables, perceived usefulness and perceived ease of use, which are hypothesized to be fundamental determinants of user acceptance. The TAM posits that users' behavioral intentions determine actual technology acceptance. Behavioral intentions are in turn influenced by the user's attitude towards technology. Davis (1989) stated that perceived usefulness and perceived ease of use are beliefs that lead to favorable attitudes and intentions to accept and use technology. However, the TAM is considerably less general than the Theory of Reasoned Action (TRA) as the former was specifically designed to apply only to computer usage behavior (Davis, Bagozzi, & Warshaw, 1989). Figure 1 shows the Technology Acceptance Model.

Although the TAM has been extensively tested and validated among end-users in the business settings, research on its application in education is limited. One reason lies in the difference between the general technology users and teachers. Teachers tend to be relatively independent and have considerable autonomy over their teaching activities, including technology choices and use. Hu, Clark, and Ma (2003) posited that because educational institutions (i.e. schools) have fundamentally different objectives compared to business organizations, teachers experience less peer competition in resources and promotions. To cope with teachers who are resistant to using computers in the classrooms, the development of positive attitudes toward computers should be considered paramount since the successful use of computers in the classroom depends to a large part on the positive attitudes of teachers (Yuen & Ma, 2002).

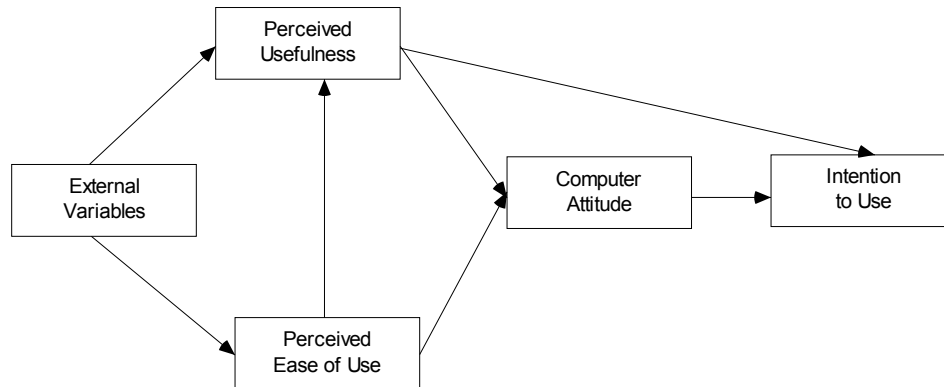


Figure 1: Technology Acceptance Model (Adapted from Davis, Bagozzi, & Warshaw, 1989)

Despite the limited applications of the TAM in educational contexts, its parsimony and predictive powers lends itself to studying the technology acceptance in educational setting. Over the years, researchers have used the TAM framework to examine users' acceptance toward various technology applications such as the Graphic User Interface (GUI) (Agarwal & Prasad, 1999), mainframe application (Dishaw & Strong, 1999), accounting applications (Jackson, Chow & Leitch, 1997), World Wide Web (Riemenschneider, Harrison, Mykytyn Jr., 2003), and computer resource centre (Taylor & Todd, 1995). More recently, researchers have used the TAM to investigate educational related issues such as student satisfaction with online learning (Drennan, Kennedy & Pisarski, 2005), student acceptance of online course companion site of a textbook (Gao, 2005) and the effect of technical support on student acceptance towards WebCT (Ngai et al. 2007). Apart from that, a recent study by Teo, Lee and Chai (2008) employed the TAM as the basic model to examine the attitudes of pre-service teachers towards the use of technology in education and found that the model predicted 42% of the variance.

This study extends the literature by applying the TAM to study the factors affecting technology acceptance among Singaporean and Malaysian pre-service teachers. The validity of the TAM has been tested across different cultures such as Switzerland, US, UK, and Arab countries (Al-Gahtani, 2001; Straub, Keil, & Brenner, 1997). A recent application of the TAM in a cross-cultural study was made by Singh, Fassott, Chao, and Hoffmann (2006) to predict the factors affecting the global consumer acceptance of international web sites among Brazilians, Germans, and Taiwanese. Thus far, no cross-cultural study has applied the TAM to study technology acceptance of pre-service teachers from Singapore and Malaysia and it would be of interest to the research community to see if the TAM can be validated in these two countries.

## Research Hypotheses

### Hypothesis for Perceived Usefulness (PU) and Perceived Ease of Use (PEU)

Perceived usefulness (PU) is defined as the degree to which a person believes that using a particular technology will enhance his or her job performance (Davis et al., 1989). People tend to use or not to use an application to the extent that they believe it will enhance their job performance (Davis et al., 1989). This includes decreasing the time for doing the job, and achieving more efficiency and accuracy. In the view of Phillips, Calantone and Lee (1994), perceived usefulness reflects the prospective users' subjective probability that applying the new technology will be beneficial to his/her personal and/or the adopting organization's well-being.

Perceived ease of use (PEU) refers to the degree to which a person believes that using a particular technology will be free of effort (Davis et al., 1989). It is possible that while users may believe that computers are useful, they may be, at the same time, too difficult to use and that the performance benefits of usage are outweighed by the effort of using the application (Davis, 1989). PEU explains the user's perception of the amount of effort required to utilize the system or the extent to which a user believes that using a particular technology will be effortless (Davis, et al., 1989). As such, it is possible that educational technology with a high level of PU is more likely to induce positive attitudes. Furthermore, the relation between PU and PEU is that PU mediates the effect of PEU on attitude (Moon & Kim,



2001). In other words, while PU has direct impacts on attitude, PEU influences attitude indirectly through PU. Therefore, the present study proposes the following hypothesis:

H1: Singaporean and Malaysian pre-service teachers' perceived usefulness of computer will be significantly influenced by their perceived ease of use of computers.

### **Computer Attitudes (CA)**

Attitudes guide behavior and attitude refers to the way an individual respond to and is disposed towards an object (Ajzen & Fishbein, 2005). This feeling or disposition may be negative or positive. The success of any initiatives to implement technology in an educational program depends strongly upon the support and attitudes of teachers involved. It has been suggested that if teachers believed or perceived computers not to be fulfilling their own or their students' needs, they are likely to resist any attempts to introduce technology into their teaching-learning process (Askar & Umay, 2001). In other words, attitudes, whether positive or negative, affect how teachers respond to the technology in an instructional setting and learning environment. This in turn affects the way students view the importance of computers in schools (Teo, 2006) and affects current and future computer usage. Yildirim (2000) further stressed that it is unlikely for teachers with negative attitudes toward computers to be able to transfer their computer skills let alone encourage the use of computers among students. Suffice to say, no matter how sophisticated and powerful the state of technology is, the extent to which it is implemented depends on teachers having a positive attitude toward it (Huang & Liaw, 2005).

A key reason for studying teachers' computer attitudes is the ability of attitudes to predict computer usage. Positive teacher attitude towards computing is critical if computers are to be effectively integrated into the school curriculum. Indeed, Woodrow (1991) emphasized that there is a need to assess teachers' attitudes as these can influence their acceptance of computers and their future use. Research has shown that a teacher's attitude towards the computer is a major predictor for future computer use (Myers & Halpin, 2002). It also predicts the need to learn computing skills that in turn will lead to computer literacy (Zhang & Espinosa, 1997). For example, Yildirim (2000) found that teachers who used computers more would tend to develop positive attitudes that promote further use of the computer in their daily teaching tasks and conduct activities that require computers to play a major role, for example, computer-mediated forums. In addition, the relationship between perceived usefulness and perceived ease of use on computer attitudes has been reported in various studies that provided evidence in support of a positive relationship among them (e.g. Cheng, Lam, & Yeung, 2006; Teo, Lee, & Chai, 2008; Yu, Ha, Choi, & Rho, 2005). This leads to two hypotheses:

H2a: Singaporean and Malaysian pre-service teachers' computer attitude will be significantly influenced by their perceived usefulness of computers.

H2b: Singaporean and Malaysian pre-service teachers' computer attitude will be significantly influenced by their perceived ease of use of computers.

### **Behavioural Intention (BI)**

Of the major difference between the TAM and TRA is the presence of behavioral intention (BI) in the TAM. The TAM implies that two behavioral beliefs, PU and PEU, have influence on an individual's intention to use technologies. In contrast to PU and PEU, which refer to process expectancy and outcome expectancy respectively (Liaw, 2002), BI leads to actual use of technologies. The validity of this claim has been demonstrated by research evidence across a variety of contexts where technology was used (e.g. Chau, 2001; Fusilier & Durlabhji, 2005). Given sufficient support to conclude the existence of a strong link between intention and actual behaviour (e.g. Mathieson, 1991; Hu, et al., 2003), BI is used as the dependent variable in this study. In addition, BI is also a practical way to measure actual use in this study. Although all pre-service teachers in this study have used technology for personal and academic reasons, most of them possess little or no experience in using technologies in the classrooms. As such it is deemed more accurate to measure respondents' intention rather than their actual use. The practice of using BI on pre-service teachers and undergraduates is widely reported in the literature (e.g. Hu, et al., 2003; Liaw & Huang, 2003). For this reason, the present study proposes these two hypotheses:

H3a: Singaporean and Malaysian pre-service teachers' behavioral intention will be significantly influenced by their perceived usefulness of computers.

H3b: Singaporean and Malaysian pre-service teachers' behavioral intention will be significantly influenced by their computer attitudes.

## Aims of this Study

The aims of this study are twofold. First, it seeks to examine if there are differences among the predictors in technology acceptance between the Singaporean and Malaysian pre-service teachers. The study also seeks to investigate the extent to which the TAM explains the intention to use technology among pre-service teachers. Figure 2 shows the research model of this study.

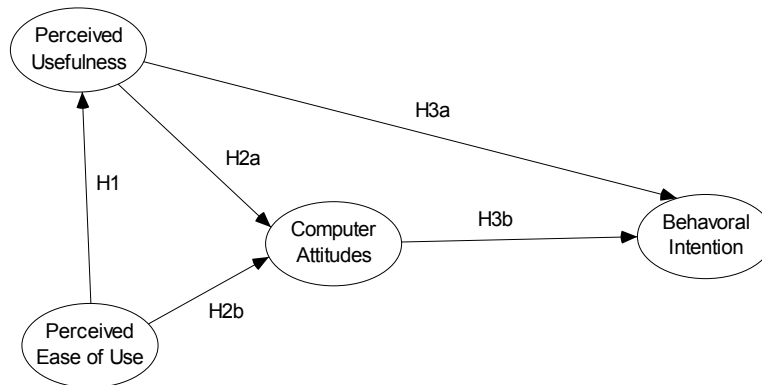


Figure 2: Research Model of this Study

## Method

### Subjects

The participants for this study comprised 250 (175 females and 75 males) and 245 (183 females and 62 females) pre-service teachers in Singapore and Malaysia respectively and all of them owned a computer at home. Participants were volunteers who were recruited in their respective country. The details of the participants are shown in Table 1.

Table 1: Demographics information of participants

Country	Age		Daily computer use		Computer experience	
	M	SD	M	SD	M	SD
Singapore	24.0	4.4	3.4	2.0	8.1	4.2
Malaysia	23.4	5.5	3.1	2.4	6.6	3.8

### Procedure

Data was collected via an online survey questionnaire. The authors in this study collected the data separately from their respective country over a period of two to three months. Participants were asked to volunteer during the study term and those who consented to participate were given a URL to access the questionnaire.

### Measures

A survey questionnaire was administered to the participants who volunteered for this study. The instrument was composed of 15 statements on PU (four items), PEU (four items), CA (five items), and BI (two items). Participants gave their opinions to each statement on a five-point Likert scale, ranging from 1 – strongly disagree to 5 – strongly agree. These items were adapted from various published sources, as shown in Table 2.

Table 2: List of Constructs and Corresponding Items

Construct	Item	
Perceived Usefulness (adapted from Davis, 1989)	PU1	Using computers will improve my work.
	PU2	Using computers will enhance my effectiveness.
	PU3	Using computers will increase my productivity.
	PU4	I find computers a useful tool in my work.
Perceived Ease of Use (adapted from Davis, 1989)	PEU1	My interaction with computers is clear and understandable.
	PEU2	I find it easy to get computers to do what I want it to do.
	PEU3	Interacting with computers does not require a lot of mental effort.
	PEU4	I find computers easy to use.
Behavioural Intention (adapted from Davis, 1989)	BI1	I will use computers in future.
	BI2	I plan to use computers often.
Computer Attitudes (adapted from Thompson et al. 1991; Compeau and Higgins, 1995)	CA1	Computers make work more interesting.
	CA2	Working with computers is fun.
	CA3	I like using computers.
	CA4	I look forward to those aspects of my job that require me to use computers.

## Results

The statistical analysis comprised two stages. The first stage examined the descriptive statistics of the measurement items and assessed the reliability and validity of the measure used in this study. The second stage tested the proposed research model and this involved assessing the contributions and significance of the manifest variables path coefficients.

### Descriptive statistics

The descriptive statistics for each construct items are shown in Table 3. All means were greater than 3.0, ranging from 3.69 to 4.46. This indicates an overall positive response to the constructs that are measured in this study. The standard deviations for all variables were less than one and this indicates that the item scores were around the mean scores. Comparing the mean scores from Singapore and Malaysia, all variables, except BI are significantly different. It was felt that this difference could be attributed to Malaysian pre-service teachers being exposed to the functions of specific IT tools (e.g. MS Office) immediately prior to completing the survey questionnaire for this survey. Their Singapore counterparts, however, had completed the questionnaire during a session in which the use of ICT in pedagogy was taught. Hence the differences in the mean values between the Singapore and Malaysian sample could be accounted by the extent to which participants had been exposed to the actual functions of ICT tools prior to data collection in this study.

Table 3: Descriptive Statistics of the Constructs

Construct	Singapore		Malaysia		t
	Mean	SD	Mean	SD	
Perceived Usefulness (PU)	4.13	.55	4.29	.61	-3.04**
Perceived Ease of Use (PEU)	3.69	.59	3.81	.63	-2.19*
Computer Attitudes (CA)	3.79	.55	3.95	.60	-3.09**
Behavioural Intention (BI)	4.46	.59	4.44	.62	.32

Note: \*\*  $p < .01$ ; \*  $p < .05$

### Factor Structure

Table 4 shows the principal component analysis of the four constructs. The total variance explained is 69.4% and factor extraction was based on the Kaiser-Guttman rule that retains principal component of eigenvalues of equal or greater than 1. This was to ensure that the factor extracted account for as much variance as the individual variable (Nunnally & Bernstein, 1994).

Table 4: Principal Component Analysis of all items

	PU	CA	BI	PEU
PU1	<b>.728</b>	.186	.180	.381
PU2	<b>.871</b>	.172	.100	.238
PU3	<b>.863</b>	.127	.137	.211
PU4	<b>.644</b>	.188	.190	.193
PEU1	.213	<b>.752</b>	.161	.168
PEU2	.278	<b>.726</b>	.087	.132
PEU3	.028	<b>.633</b>	-.023	.170
PEU4	.107	<b>.805</b>	.120	.204
BI1	.213	.096	<b>.835</b>	.148
BI2	.159	.114	<b>.855</b>	.119
CA1	.358	.077	.174	<b>.733</b>
CA2	.297	.250	.168	<b>.709</b>
CA3	.205	.242	.193	<b>.775</b>
CA4	.168	.266	-.043	<b>.745</b>
Total Var. Expl.	2.940	2.478	1.671	2.631
% Var. Expl.	21.00	17.70	11.94	18.79

Rotation Method: Varimax with Kaiser Normalization.

### Convergent validity

Fornell and Larcker (1981) proposed three procedures to assess for convergent validity of a set of measurement items in relation to their corresponding constructs. These are (1) item reliability of each measure, (2) composite reliability of each construct and (3) the average variance extracted.

The item reliability of an item was assessed by its factor loading onto the underlying construct. Hair et al. (2006) suggested that an item is significant if its factor loading is greater than 0.50. As shown in Table 4, the eigenvalues of all constructs exceeded 1.00 and the percent of cumulative variance explained of these five constructs was 69.4%. The factor loadings of all the items in the measure ranged from 0.64 to 0.86. This exceeds the threshold set by Hair et al. (2006) and demonstrates convergent validity at the item level.

The composite reliability of each construct was assessed using Cronbach's alpha. Robinson, Shaver and Wrightsman (1991) and DeVellis (2003) suggested that an alpha value of .70 should be considered acceptable. As shown in Table 5, the reliabilities of all the constructs are between .73 and .88, well within the range suggested by Robinson et al. (1991) and DeVellis (2003).

The final indicator of convergent validity, average variance extracted, is a more conservative test of convergent validity (Fornell & Larcker, 1981). It measures the amount of variance captured by the construct in relation to the amount of variance attributable to measurement error. Convergent validity is judged to be adequate when average variance extracted equals or exceeds 0.50 (i.e. when the variance captured by the construct exceeds the variance due to measurement error). As shown in Table 5, the convergent validity for the proposed constructs of the research model is adequate.

Table 5: Construct Reliability and Average Variance Extracted

Construct	Cronbach's $\alpha$	AVE
Perceived Usefulness (PU)	.88	.61
Perceived Ease of Use (PEU)	.76	.54
Computer Attitude (CA)	.75	.55
Behavioral Intention (BI)	.73	.71

AVE: Average Variance Extracted. This is computed by squaring the sum of factor loading divided by number of factors of the underlying construct.

## Discriminant validity

Discriminant validity is assessed to measure the extent to which constructs are different. At the item level, Barclay, Higgins and Thompson (1995) suggested that discriminant validity is present when an item correlates more highly with items in the construct it intends to measure more than items belonging to other constructs. In this study, an acceptable level of discriminant validity at the item level was found (Appendix 1).

At the construct level, discriminant validity is considered adequate when the variance shared between a construct and any other construct in the model is less than the variance that construct shares with its measures (Fornell, Tellis & Zinkham, 1982). The variance shared by any two constructs is obtained by squaring the correlation between the two constructs. The variance shared between a construct and its measures corresponds to average variance extracted. Discriminant validity was assessed by comparing the square root of the average variance extracted for a given construct with the correlations between that construct and all other constructs. Table 6 shows the correlation matrix for the constructs. The diagonal elements have been replaced by the square roots of the average variance extracted. For discriminant validity to be judged adequate, these diagonal elements should be greater than the off-diagonal elements in the corresponding rows and columns. Discriminant validity appears satisfactory at the construct level in the case of all constructs. This indicates that each construct shared more variance with its items than it does with other constructs. Having achieved discriminant validity at both the item and construct levels, the constructs in the proposed research model are deemed to be adequate.

Table 6: Inter-construct Correlation Matrix\*

	PU	PEU	CA	BI
PU	(.78)			
PEU	.44	(.73)		
CA	.62	.51	(.74)	
BI	.43	.29	.36	(.84)

Notes: \* =  $p < .01$ ; Diagonal in parentheses: square root of average variance extracted from observed variables (items); Off-diagonal: correlations between constructs

## Model Fit

### Test of the proposed model

Structural equation modeling (SEM) was performed to test the fit between the research model (Figure 2) and the obtained data. This technique is chosen for its ability to examine a series of dependence relationships simultaneously, especially where there are direct and indirect effects among the constructs within the model (Hair et al., 2006). In this study, AMOS 7.0 (Arbuckle, 2006) was used and the SEM estimation procedure is maximum likelihood estimation. In the SEM, the sample size plays an influential role in the reliability of the result. Bollen (1989) recommended a minimum sample size of 100 while Anderson and Gerbing (1988) recommended 200. A recent proposal by Hair et al. (2006) indicated that any study with five or fewer constructs, each with more than three items, and high item communality with .60 and higher, can adequately be estimated with sample size of 150. In this study, the sample size is 485 (250+245) and this was considered adequate on the basis of recommendation from research. Although Behavioral Intention (BI) comprised only two items, it has an acceptable alpha coefficient and AVE, indicating an acceptable level of convergent validity. On this basis, the authors decided that this is an acceptable minor deviation from Hair et al. (2006) recommendations for this study.

In using SEM, it is a common practice to use a variety of indices to measure model fit (Kline, 2005). In addition to the ratio of the  $\chi^2$  statistic to its degree of freedom, with a value less than 5 indicating acceptable fit, researchers recommended a handful of fit indices to assess model fit (e.g. Kline, 2005). These are the Goodness of Fit (GFI), Normed Fit Index (NFI), Standardized Root Mean Residual (SRMR), and the Comparative Fit Index (CFI). Table 7 shows the level of acceptable fit and the fit indices for the proposed research model in this study. Except for the  $\chi^2$ , all values satisfied the recommended level of acceptable fit. In the case of the  $\chi^2$ , it has been found to be too sensitive to sample size differences, especially for cases in which the sample size exceeds 200. Hair et al., (2006) noted that, as the sample size increases, there is a great tendency for the  $\chi^2$  to indicate significant differences. Therefore, this anomaly is assumed to be applicable in the present study with a sample of 495. However, the results of the  $\chi^2 / df$

value in the present study is well within the recommended  $\chi^2 / df < 5$ . Table 7 shows the model fit for the overall model and separately for the Singaporean and Malaysian samples. There is a good fit for both models.

Table 7: Fit Indices of the Proposed Research Model

Fit Index	Recommended Level of Fit	Overall Model	Singapore	Malaysia
$\chi^2$	n.s at $p < .05$	230.9, s. at $p < .05$	149.9, s, $p < .05$	163.0, s, $p < .05$
$\chi^2 / df$	$< 5$	2.75	1.81	1.94
GFI	$> 0.90$	0.94	0.93	0.92
NFI	$> 0.90$	0.93	0.90	0.93
SRMR	$< 0.05$	0.04	0.05	0.04
RMSEA	$< 0.05$	0.05	0.05	0.06
CFI	$> 0.90$	0.96	0.95	0.96

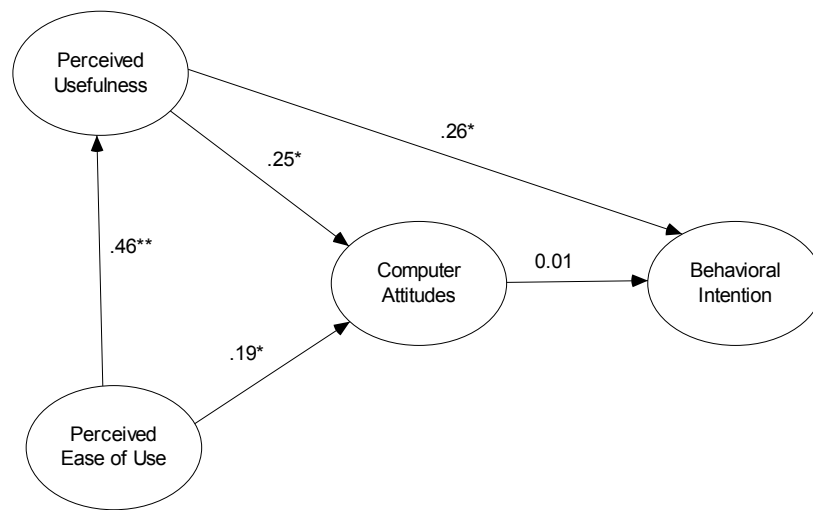


Figure 3: Result of the Singaporean Sample

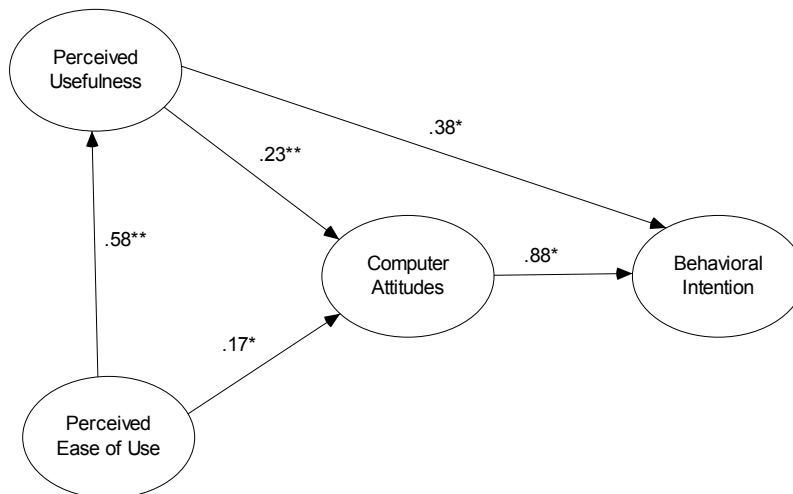


Figure 4: Result of the Malaysian Sample

Figures 3 and 4 show the resulting path coefficients of the proposed research model for the Singaporean and Malaysian samples respectively. Except for 3b, all hypotheses were supported by the data. The results show that perceived ease of use significantly influenced perceived usefulness ( $\beta = 0.54$ ,  $p < 0.01$ ), supporting hypothesis H1.

Computer attitudes was influenced by perceived usefulness ( $\beta = 0.25$ ,  $p < 0.01$ ) and ease of use ( $\beta = 0.17$ ,  $p < 0.01$ ), supporting hypotheses H2a and H2b. Behavioral intention was found to be influenced by perceived usefulness ( $\beta = 0.29$ ,  $p < 0.01$ ) and computer attitudes ( $\beta = 0.56$ ,  $p < 0.05$ ), thus supporting hypotheses H3a and H3b. However, hypothesis H3b does not hold for the Singaporean sample (0.01, n.s.) although this is true for the Malaysian sample ( $\beta = 0.88$ ,  $p < 0.01$ ). Three endogenous variables were tested in the overall model. Perceived usefulness was found to be significantly determined by perceived ease of use, resulting in an  $R^2$  of 0.27. This means that perceived ease of use explained 27% percent of the variance in perceived usefulness. Computer attitudes was significantly determined by perceived usefulness and perceived ease of use and the percent of variance explained was 59% ( $R^2 = 0.59$ ). The dependent variable, behavioral intention was significantly determined by perceived usefulness, and perceived ease of use, and computer attitude resulting in an  $R^2 = 0.29$ . This means that the combined effects of perceived usefulness, perceived of ease, and computer attitude explained 29% of the variance of behavioral intention. A summary of the hypotheses testing results is shown in Table 8.

*Table 8: Hypothesis Testing Results*

Hypotheses	Path	Overall Model	Singapore	Malaysia	Results
H1	PEU $\rightarrow$ PU	.54**	.46**	.59**	Supported
H2a	PU $\rightarrow$ CA	.25**	.25*	.23**	Supported
H2b	PEU $\rightarrow$ CA	.17**	.18*	.17*	Supported
H3a	PU $\rightarrow$ BI	.29**	.26*	.39**	Supported
H3b	CA $\rightarrow$ BI	.56*	.01	.88*	Not Supported

\*  $p < .05$ ; \*\*  $p < .001$

## Discussion

This study aims to explore pre-service teachers' intention to use technology and found perceived usefulness, perceived ease of use, computer attitudes to be significant determinants. On whether there are differences among the predictors of technology acceptance between the Singaporean and Malaysian pre-service teachers and the extent to which the TAM explains technology acceptance among pre-service teachers, this study shows that all variables except BI were significantly different. This suggests that while the predictor variables (PU, PEU, and CA) may interact differently for each of the two samples, there was no difference in the way the dependent variable (BI) had been treated by the participants. This finding indicates that both Singaporean and Malaysian pre-service teachers would most likely integrate technology into their teaching-learning process or administrative work.

All hypotheses were supported except for H3b. While the path from CA to BI is significant for the Malaysian sample, this is not so for the Singaporean sample. The relationship between CA and BI has been found to be influenced by a user's volition. Venkatesh et al. (2003) suggested that computer attitude was a significant predictor of the intention to use in situations where the use of technology was volitional. It was possible that the Singaporean sample had perceived that the use technology was an obligation. This could be due to the fact that the pre-service teachers in this study had been strongly encouraged to use computers by having to (1) attend compulsory modules in using ICT for instructional purposes; (2) use computers for assessment purposes, and (3) rely on e-learning portals for information and course administration. It is possible that such experiences have portrayed an impression of a 'mandate' to the participants in this study. On the other hand, the Malaysian sample was trained in a slightly different learning environment than that of their Singaporean counterparts. They were enrolled in a discrete ICT course where they were exposed to new ICT knowledge and skills. In terms of their assessment, almost 50% of their exercises completed during their laboratory sessions were not graded. It was possible that, in such learning environment, the Malaysian sample felt that they have more control over the use of technology, relative to their Singaporean counterparts.

This study contributes to the literature by validating the TAM in two different cultures. Previous studies have highlighted the need to extend the TAM to different contexts to be validated in different cultures to enhance its generalizability (Yoo & Donthu, 2001; Zeithaml, Parasuraman & Malhotra, 2000). The present study is an attempt to build on previous research and validate the TAM in two countries. However, it is important to note that the link between CA and BI is not significant for the Singapore sample, hence H3b was not supported. As such, the proposed

model in this study can at best be only considered as partially validated. Further research is needed to ascertain the relationship between CA and BI.

Although the TAM has been validated in the USA, UK, Switzerland, Germany, Brazil, some Arab countries, it has been tested in only a few Asian countries and one of these is Taiwan (Singh et al., 2006). Furthermore, many studies applying the TAM selected their participants from the same country and it is argumentative whether culture did mediate acceptance in such cases since all participants share the same culture. This study included participants that were taken from two different countries. The samples in this study came from Singapore and Malaysia, both of whose population comprises people from different ethnic groups which includes the Chinese, Malays, Indians, and Eurasians (descendants of the early European settlers). Most of these ethnic groups are represented by the participants in this study. This is in contrast to other cross-cultural studies that comprised mainly mono-cultural samples (e.g. Chai & Pavlou, 2004; Straub et al. 1997). It is reasonable to expect that using participants from diverse races contribute to a greater understanding of the role of TAM as a valid model to explain technology acceptance. However, it should also be noted that more than 70% of the Singaporean pre-service teachers are Chinese whereas the Malaysian pre-service teachers are predominantly Malays.

This study provides a cross-cultural validation of TAM and in a different context. Many cross-cultural validation of TAM have been conducted in non-educational contexts. Recent examples include the use of IT in an organizational setting (e.g. Heijden, 2003) or consumer setting (e.g. Singh et al., 2006). As discussed in the literature section of this paper, teachers play a key role in mediating technology acceptance in the schools, it is important to understand the determinants of their behavioural intentions. The results of this study show that Singaporean and Malaysian pre-service teachers do not differ in their behavioural intentions with regards to technology acceptance. Although these variables have been found to be significant in this study, their interactions with other external factors are important within the TAM framework and warrant further attention.

While it is important to ensure that teachers are able to use technology in the curriculum, the groundwork must be laid at the pre-service teacher's level. In the course of their training, pre-service teachers should be provided with tools and experiences that will be used regularly in their future job as a teacher as teaching has become a highly complex activity. The effective use of technology enables teachers facilitate and adjust their instructional strategies to optimise students' learning. This is likely to happen if teachers' attitudes toward computers are considered (Yuen, Law & Chan, 1999).

On the part of teacher educators, there is a need to understand the dimensions that influence pre-service teachers' attitudes towards computers to design teacher-training curricula that will prepare teachers to face the challenges in the information age (Fisher, 2000). Since attitude has been shown to predict usage, a good understanding of pre-service teachers' computer attitudes will shed light on future computer use by in-service (trained) teachers. Within the framework of TAM, attitude has been used as a dependent variable to study possible linkages to computer usage (Yang & Yoo, 2004).

Future research could be conducted to collect data from practicing teachers to compare the results with those from pre-service teachers. Longitudinal studies may be designed to trace the stages of attitudinal changes experienced by pre-service teachers when they become practicing teachers. Finally, it is useful to examine whether there are discrepancies between self-reports and actual practice and, if these exist, to identify the factors that explain the gap.

## **Limitations**

This study has several limitations. Firstly, pre-service teachers were used as participants and their views may differ from the practicing teachers. This is an important consideration given that the consequences for inadequate technology use are borne by practicing teachers and their students. In addition, practicing teachers are more likely than pre-service teachers to be exposed to demands in relation to their use of computers from within and outside their professional environments. Pre-service teachers may not fully appreciate the demands and stress involved in the use of computers in a real-life school setting despite having undergone school experiences and teaching practice in the course of their training. Such mandatory versus voluntary use of computers has been found to influence teachers' views of technology usage as a whole (e.g., Legris et al. 2003).



Secondly, despite careful attention given to the methodology, the study data were self-reported. This method of data collection to study attitude raises the possibility of common method variance, a situation in which variations among the scores are due to the method of data collection instead of the intended constructs. This may give an inflated association between constructs (Kline, Sulksy & Rever-Moriyama, 2000).

## Future Research

This study did not account for the participants' computer experience and computer usage situation. As such, the variations introduced by not including these two variables may have contributed to the lack of support for H3b. Such subject-related variables should be included in future research.

Methodologically, future research on computer attitude should employ a multi-trait multi-method matrix (MTMM), an approach that has been suggested to minimise common method variance that are associated with a single method of data collection. In terms of statistical treatment, we suggest on uses of SEM to measure structural invariance of the TAM to test of its validity across culture. Other possibilities include:

- a. Cross-cultural study of TAM among mono-and multi-cultural samples;
- b. Extend TAM to include other variables of interest to the education community;
- c. Study how additional variables may cause TAM to vary in their importance at various levels of technology acceptance;
- d. Study how competing models such as TRA and TPB (which focuses on explaining attitude) compare with TAM in terms of predicting usage of technology;
- e. Use TAM to examine acceptance of various forms of technologies and contexts to explore its validity.

## Conclusion

This study aims to examine possible differences in technology acceptance between Singaporean and Malaysian pre-service teachers. Results show no significant difference in BI although there were significant differences in the exogenous variables (PU, PEU and CA). The results also show that the causal relationships among the TAM constructs (PU, PEU, and CA) and the outcome variable (BI) were as predicted in previous literature. As such, there is evidence from this study to support the applicability in explaining technology acceptance of pre-service teachers in Singapore and Malaysia.

## References

- Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences*, 30(2), 361-391.
- Ajzen, I., & Fishbein, M. (1980) *Understanding attitudes and predicting social behavior*, Englewood Cliffs, NJ: Prentice-Hall.
- Ajzen, I., & Fishbein, M. (2005) The influence of attitudes on behavior, in: D. Albarracín, B. T. Johnson & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change: basic principles*. Mahwah, NJ: Erlbaum.
- Al-Gahtani, Said S. (2001). The applicability of TAM outside North America: An empirical test in the United Kingdom. *Information Resources Management Journal*, 14(3), 37-46.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423.
- Arbuckle, J. L. (2006). *AMOS (version 6.0) [Computer software]*, Chicago: SmallWaters.
- Askar, P., & Umay, A. (2001). Pre-service elementary mathematics teachers' computer self-efficacy, attitudes towards computers, and their perceptions of computer-enriched learning environments. In *Proceedings of Society for Information Technology and Teacher Education International Conference*, Chesapeake, VA: AACE, 2262-2263.

- Barclay, D., Higgins, C., & Thompson, R. (1995) The Partial Least Squares (PLS) approach to causal modeling: Personal computer adoption and use as an illustration. *Technology Studies*, 2(2), 285–309.
- Becker, H. (2001). How are teachers using computers in instruction? *Paper presented at the 2001 Meetings of the American Educational Research Association*, Retrieved 19 September 2008 from: [http://www.crito.uci.edu/tlc/findings/conferences-pdf/how\\_are\\_teachers\\_using.pdf](http://www.crito.uci.edu/tlc/findings/conferences-pdf/how_are_teachers_using.pdf).
- Bereiter, C., & Scardamalia, M. (2006) Education for the knowledge age. In P. A. Alexander, & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2<sup>nd</sup> Ed.), Mahwah, NJ : Lawrence Erlbaum, 695-713.
- Bollen, K. A. (1989). *Structural equations with latent variables*, New York: John Wiley.
- Burns, M. (2002). From black to white to color: technology, professional development and changing practice. *T.H.E. Journal online*, Retrieved September 18, 2008 from: <http://thejournal.com/articles/16020>.
- Chai, L., & Pavlou, P. A. (2004). From ‘ancient’ to ‘modern’: A cross-cultural investigation of electronic commerce adoption in Greece and the United States. *Journal of Enterprise Information Management*, 17(6), 416-23.
- Chau, P. Y. K. (2001). Influence of computer attitude and self-efficacy on IT usage behaviour. *Journal of End-User Computing*, 13(1), 26-33.
- Cheng, T. C. E., Lam, D. Y. C., & Yeung, A. C. L. (2006), Adoption of internet banking: An empirical study in Hong Kong, *Decision Support Systems*, 42(3), 1558-1572.
- Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189-211.
- Davis, F. D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989) User acceptance of computer technology A comparison of two theoretical models. *Management Science*, 35(8), 928-1003.
- DeVellis, R. F. (2003). *Scale development: theory and applications*. (2<sup>nd</sup> Ed.), Newbury Park, CA: Sage.
- Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task-technology fit constructs. *Information & Management*, 36(1), 9-21.
- Drennan, J., Kennedy, J., & Pisarski, A. (2005). Factors affecting student attitudes toward flexible online learning in management education. *The Journal of Educational Research*, 98(6), 331-340.
- Fisher, M., (2000) Computer skills of initial teacher education students. *Journal of Information Technology for Teacher Education*, 9(1), 109-123.
- Fornell, C., & Larcker, D. F. (1981) Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 48, 39-50.
- Fornell, C., Tellis, G. J., & Zinkhan, G. M. (1982) *Validity assessment: A structural equations approach using partial least squares*. *AMA Educator's Proceedings*, Chicago, 405-409.
- Fusilier, M., & Durlabhji, S. (2005). An exploration of student internet use in India. *Campus-Wide Information Systems*, 22(4), 233-246.
- Gao, Y. (2005). Applying the technology acceptance model (TAM) to educational hypermedia: A field study. *Journal of Educational Multimedia and Hypermedia*, 14(3), 237-247.
- Gong, M., Xu, Y., & Yu, Y. (2004). An Enhanced Technology Acceptance Model for Web-Based Learning. *Journal of Information Systems Education*, 15(4), 365-374.
- Hair, J. F. Jr., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006) *Multivariate data analysis* (6<sup>th</sup> Ed.), New Jersey: Prentice-Hall.
- Heijden, H. V. D. (2003). Factors influencing the usage of websites: The case of a generic portal in The Netherlands. *Information & Management*, 40(6), 541-549.
- Hu, P. J., Chau, P. Y. K., Liu-Sheng, O. R., & Yan-Tam, K. (1999). Examining the technology acceptance model using physician acceptance of telemedicine technology. *Journal of Management Information Systems*, 16(2), 91-112.

- Hu, P. J., Clark, T. H. K., & Ma, W. W. (2003) Examining technology acceptance by school teachers: A longitudinal study. *Information & Management*, 41(2), 227-241.
- Huang, H. M., & Liaw, S. S. (2005). Exploring user's attitudes and intentions toward the web as a survey tool. *Computers in Human Behavior*, 21(5), 729-743.
- Jackson, C.M., Chow, S., & Leitch, R.A. (1997). Toward an understanding of the behavioral intention to use an information system. *Decision Sciences*, 28(2), 357-89.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2<sup>nd</sup> Ed.), New York: Guilford Press.
- Kline, T. J. B., Sulksy, L. M., & Rever-Moriyama, S. D. (2000). Common method variance and specification errors: A practical approach to detection. *The Journal of Psychology*, 134(4), 401-421.
- Lee, A. C. K. (2003). Undergraduate students' gender differences in IT skills and attitudes. *Journal of Computer Assisted Learning*, 19, 488-500.
- Legris, P., Ingham, J., & Colletette, P. (2003). Who do people use information technology? A critical review of the technology acceptance model. *Information & Management*, 40, 191-204.
- Liaw, S.S. (2002). Understanding user perceptions of world-wide web environments. *Journal of Computer Assisted Learning*, 18(2), 137-148.
- Liaw, S. S., & Huang, H. M. (2003). An investigation of users' attitudes toward search engines as an information retrieval tool. *Computers in Human Behavior*, 19, 751-765.
- Mathieson, K. (1991). Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research*, 2(3), 173-191.
- McCoy, S., Galletta, D., & King, W. (2007). Applying TAM across cultures: the need for caution. *European Journal of Information System*, 16, 81-90.
- Moon, J., & Kim, Y. (2001). Extending the TAM for a world-wide-web context. *Information & Management*, 38(4), 217-230.
- Myers, J. M., & Halpin, R. (2002) Teachers' attitudes and use of multimedia technology in the classroom: Constructivist-Based professional development training for school districts. *Journal of Computing in Teacher Education*, 18(4), 133-140.
- Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers and Education*, 48, 250-267.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Phillips, L. A., Calantone, R., & Lee, M. T. (1994). International technology adoption: Behavior structure, demand certainty and culture. *Journal of Business & Industrial Marketing*, 9(2), 16-28.
- Riemenschneider, C. K., Harrison, D. A., & Mykytn, P. P. Jr. (2003) Understanding IT adoption decisions in small business: Integrating current theories. *Information & Management*, 40(4), 269-285.
- Robinson, J.P., Shaver, P.R., & Wrightsman, L.S. (1991). *Measures of personality and social psychological attitudes*, San Diego, CA: Academic Press.
- Singh, N., Fassott, G., Chao, M. C. H., & Hoffmann, J. A. (2006). Understanding international Web site usage: A cross-national study of German, Brazilian, and Taiwanese online consumers. *International Marketing Review*, 23, 83-97.
- Sivo, S., & Pan, C. (2005). Undergraduate engineering and psychology students' use of a course management system: A factorial invariance study of user's characteristics and attitudes. *Journal of Technology Studies*, 31(2), 94-103.
- Smarkola, C. (2007). Technology acceptance predictors among student teachers and experienced classroom teachers. *Journal of Educational Computing Research*, 31(1), 65-82.
- Straub, D., Keil, M., & Brenner, W. (1997). Testing the technology acceptance model across cultures: A three country study. *Information & Management*, 33(1), 1-11.

- Taylor, S., & Todd, P. A. (1995) Understanding information technology usage: A test of competing models. *Information Systems Research*, 6(2), 144-176.
- Teo, T. (2006) Attitudes toward computers: A study of post-secondary students in Singapore. *Interactive Learning Environments*, 14(1), 17-24.
- Teo, T., Lee, C. B., & Chai, C. S. (2008). Understanding pre-service teachers' computer attitudes: Applying and extending the Technology Acceptance Model (TAM). *Journal of Computer Assisted Learning*, 24(2), 128-143.
- Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, 15(1), 124-143.
- Thong, J. Y. L., Hong, W. Y., & Tam, K. Y. (2002). Understanding user acceptance of digital libraries: What are the roles of interface characteristics, organizational context, and individual differences? *International Journal of Human-Computer Studies*, 57(3), 215-242.
- Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27, 425-478.
- Woodrow, J. E. J. (1991). A comparison of four computer attitude scales. *Journal of Educational Computing Research*, 7(2), 165-187.
- Wozney, L., Venkatesh, V., & Abrami, P. C. (2006). Implementing computer technologies: Teachers' perceptions and practices. *Journal of Technology and Teacher Education*, 14(1), 173-207.
- Yang, H. D., & Yoo, Y. (2004). It's all about attitude: Revisiting the technology acceptance model. *Decision Support Systems*, 38(1), 19-31.
- Yildirim, S. (2000) Effects of an educational computing course on pre-service and inservice teachers: A discussion and analysis of attitudes and use. *Journal of Research on Computing in Education*, 32(4), 479-495.
- Yoo, B., & Donthu, N. (2001). Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research*, 52, 1-14.
- Yu, J., Ha, I., Choi, M., & Rho, J. (2005). Extending the TAM for t-commerce. *Information & Management*, 42(7), 965-976.
- Yuen, H.K., Law, N., & Chan, H. (1999). Improving IT training for serving teachers through evaluation. In G. Cumming, T. Okamoto & L. Gomez (Eds.), *Advanced Research in Computers and Communications in Education (Vol. 2)*, Amsterdam: IOS Press, 441-448.
- Yuen, A., & Ma, W. (2002). Gender Differences in Teacher Computer Acceptance. *Journal of Technology and Teacher Education*, 10 (3), 365-382.
- Zhang, Y., & Espinoza, S. (1997) Affiliations of computer self-efficacy and attitudes with need for learning computer skills. *Journal of Educational Computing Research*, 17, 371-383.
- Zeithaml, V.A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: a critical review of extant knowledge. *Journal of the Academy of Marketing Science*, 30, 362-375.

## Appendix 1

### Inter-correlations among the items

	PU1	PU2	PU3	PU4	PEU1	PEU2	PEU3	PEU4	BI1	BI2	CA1	CA2	CA3	CA4	CA5
PU1	1	.751(**)	.688(**)	.477(**)	.391(**)	.395(**)	.238(**)	.306(**)	.384(**)	.333(**)	.559(**)	.562(**)	.513(**)	.414(**)	.092(*)
PU2	.751(**)	1	.801(**)	.556(**)	.356(**)	.365(**)	.185(**)	.331(**)	.329(**)	.291(**)	.505(**)	.466(**)	.423(**)	.393(**)	.093(*)
PU3	.688(**)	.801(**)	1	.545(**)	.344(**)	.356(**)	.167(**)	.246(**)	.346(**)	.308(**)	.482(**)	.434(**)	.430(**)	.356(**)	.142(**)
PU4	.477(**)	.556(**)	.545(**)	1	.283(**)	.316(**)	.196(**)	.297(**)	.314(**)	.304(**)	.382(**)	.406(**)	.394(**)	.324(**)	.037
PEU1	.391(**)	.356(**)	.344(**)	.283(**)	1	.545(**)	.345(**)	.609(**)	.277(**)	.223(**)	.332(**)	.429(**)	.388(**)	.299(**)	.119(**)
PEU2	.395(**)	.365(**)	.356(**)	.316(**)	.545(**)	1	.326(**)	.546(**)	.195(**)	.247(**)	.297(**)	.392(**)	.338(**)	.357(**)	.097(*)
PEU3	.238(**)	.185(**)	.167(**)	.196(**)	.345(**)	.326(**)	1	.390(**)	.146(**)	.085	.191(**)	.202(**)	.275(**)	.301(**)	.062
PEU4	.306(**)	.331(**)	.246(**)	.297(**)	.609(**)	.546(**)	.390(**)	1	.207(**)	.232(**)	.274(**)	.411(**)	.396(**)	.358(**)	.072
BI1	.384(**)	.329(**)	.346(**)	.314(**)	.277(**)	.195(**)	.146(**)	.207(**)	1	.577(**)	.349(**)	.311(**)	.310(**)	.161(**)	.014
BI2	.333(**)	.291(**)	.308(**)	.304(**)	.223(**)	.247(**)	.085	.232(**)	.577(**)	1	.262(**)	.273(**)	.319(**)	.180(**)	.014
CA1	.559(**)	.505(**)	.482(**)	.382(**)	.332(**)	.297(**)	.191(**)	.274(**)	.349(**)	.262(**)	1	.645(**)	.573(**)	.494(**)	.176(**)
CA2	.562(**)	.466(**)	.434(**)	.406(**)	.429(**)	.392(**)	.202(**)	.411(**)	.311(**)	.273(**)	.645(**)	1	.620(**)	.477(**)	.214(**)
CA3	.513(**)	.423(**)	.430(**)	.394(**)	.388(**)	.338(**)	.275(**)	.396(**)	.310(**)	.319(**)	.573(**)	.620(**)	1	.594(**)	.202(**)
CA4	.414(**)	.393(**)	.356(**)	.324(**)	.299(**)	.357(**)	.301(**)	.358(**)	.161(**)	.180(**)	.494(**)	.477(**)	.594(**)	1	.228(**)
CA5	.092(*)	.093(*)	.142(**)	.037	.119(**)	.097(*)	.062	.072	.014	.014	.176(**)	.214(**)	.202(**)	.228(**)	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

# Educational Technology at a Crossroads: Examining the Development of the Academic Field in Canada

**Rocci Luppardini**

Communication Department, University of Ottawa, 554 King Edward RM. 310, Ottawa, ON K1N6N5, Canada // Tel: +1 (613)562-5800 ext 2519 // [rluppardini@uottawa.ca](mailto:rluppardini@uottawa.ca)

## ABSTRACT

The purpose of this qualitative research study is to gauge the current state of the academic developments in Educational Technology in order to identify pivotal issues and offer suggestions for future planning in Canada. This article explores the professional literature and the views of 25 senior faculty members from the twelve Canadian universities which offer or offered graduate degree programs in Educational Technology (and Distance Education). Findings from the study revealed one main category (challenges) connected to the academic field of Educational Technology in Canada with seven key sub-categories, namely, identity, standardization and professionalization, university politics, external influence, competition, funding, and teaching and learning. A synthesis of information sources is provided to delineate major patterns and generate new theory to help guide strategic program planning and evaluation. Recommendations suggest that greater attention should be invested in partnering and identity branding activities within the field to help leverage program success.

## Keywords

Educational Technology, Distance Education, Technology, Higher Education, Curriculum

## Introduction

James Finn once stated, “The educational future will belong to those who can grasp the significance of instructional technology” (Finn, 1964). Finn’s visionary remark embodies the promise and potential of Educational Technology in the 21<sup>st</sup> century as it did over forty years ago. Early efforts to conceptualize Educational Technology and its development originated in the maturation of the audio-visual movement in education and instructional training programs beginning in the First World War (Reiser, 1987).

In the 1970s and 1980s, the evolving academic field of Educational Technology was recognized as complex and multi-faceted. Davies (1978) ascribed three levels of evolution in the field of Educational Technology corresponding to a tool level (Educational Technology I), systematic level (Educational Technology II), and a systemic level (Educational Technology III). By the end of the 1980s, Beckwith (1988) described the field at the systematic level (Educational Technology II) state of evolution (Beckwith, 1988). Beckwith (1988) envisioned the future evolution of the field where the creation of unified approach at the systemic level (Educational Technology III) would transform learning while operating within the confines of a systematic level (Educational Technology II).

In Beckwith’s view, the field of Educational Technology appeared to be focused on the survival of the systematic level (Educational Technology II) at the expense of the realization of a unified approach at the systemic level of evolution (Educational Technology III). According to Pals and Plomp (1989), the tools approach is centered on physical media developed to assist in the teaching/learning process, the systematic focused on processes, used for developing, designing and evaluating instruction, and the systemic approach has a philosophical and holistic orientation for contextualizing, analyzing and solving problems through a consideration of as many facets and their interrelationship as possible. The holist commitment to the whole as being greater than the sum of the parts entails that individuals are part of the system in which they observe and, therefore, cannot achieve complete knowledge of the whole. The systemic approach is aligned with postmodernist perspectives and new paradigms in Educational Technology (Hynka, 1994, Luppardini, 2005; Yeaman, Hynka, Anderson, Damarin, & Muffoletto, 1996).

Ongoing efforts to conceptualize the evolving field of Educational Technology have spurred debate and given rise to a growing body of literature focusing on the state of the field, key challenges, and possible future developments (Beckwith, 1988; Clark, 1978; Heinich, 1984; Hynka & Nelson, 1985; Mitchell, 1989; Silber, 1970, 1978; Torkelson, 1987). Also, an entire issue of the *Canadian Journal of Educational Communication* (1989) was dedicated to debates covering the future of the field. This is partly due to the nature of applied fields like Educational

Technology where multiple knowledge bases are employed (Luppigini, 2005). Luppigini (2005) noted that the development of new knowledge in this applied field causes shifts in thinking and introduces change in multiple knowledge bases, this, multiplying change and compounding difficulties in reaching any unified perspective. However, for the purpose of this study, the definition of Educational Technology follows Luppigini's (2005) conceptual study of the field:

Educational Technology is a goal oriented problem-solving systems approach utilizing tools, techniques, theories, and methods from multiple knowledge domains, to: (1) design, develop, and evaluate, human and mechanical resources efficiently and effectively in order to facilitate and leverage all aspects of learning, and (2) guide change agency and transformation of educational systems and practices in order to contribute to influencing change in society.

Perhaps the most significant recent development in the academic field of Educational Technology in Canada concerns converging aspects of Educational Technology and Distance Education in Canada as indicated by the growth of Distance Education programs in Canada and the amalgamation of Canada's premier Educational Technology and Distance Education professional Associations, the Canadian Association of Distance Education (CADE) and Association for Media and Technology in Education in Canada (AMTEC), into one unified association called the Canadian Network for Innovation in Education (AMTEC, 2008). In line with this development, the present study also includes distance education and educational technology degree programs in its examination of the academic field of Educational Technology.

This study attempts to understand how the academic field of educational technology has developed in Canada since the creation of the first graduate degree program in 1968 and, from that understanding, to contribute insights about its evolution and current state. Building on the discussion of established approaches and early challenges to Educational Technology (Davies, 1978; Hlynka and Nelson, 1985), this researcher is curious about current challenges and possible future development within the academic field of Educational Technology and Distance Education. Specifically, this study is divided into two sections: (1) it documents the history of the academic field of Educational Technology in Canada and assesses its current state from the viewpoints of senior faculty who contributed to creating and discontinuing degree programs; and (2) it generates grounded theory to help guide academic program decision-making. To this end, the study draws on qualitative research methods.

## **Method**

### **Research Design**

This qualitative research is a historical study that draws on the grounded theory approach of Strauss and Corbin (1998) to guide data collection and analysis. 'Grounded theories, because they are drawn from data, are likely to offer insight, enhance understanding, and provide a meaningful guide to action' (Strauss & Corbin, 1998, p. 12). The following research question guided this study: How has the academic field of educational technology in Canada arrived at the point it is at now? To answer this research question, both interview data were required from individuals who have an 'insider' perspective on aspects of the academic field of educational technology and documents that are not assessable in the public realm (Trice & Beyer, 1993).

### **Sample and Participants**

Sampling of educational technology programs was determined by the following criteria: (1) degree programs acknowledged by the main national association for educational technology (AMTEC), (2) list of degree programs published in the annual Athabasca University Directory of Canadian university programs, (3) degree titles and similar titles from Canadian universities acknowledged by AMTEC. Since 1968, 12 Canadian universities have created degree programs in educational technology under several titles: Concordia University (Educational Technology), Acadia University (Learning and Technology), Memorial University (Information Technology), University of Calgary (Educational Technology), Athabasca University (Distance Education), University of Alberta (Instructional Technology), University of Saskatchewan and (Educational Communications and Technology), Laval University (Technologie éducative), Tele-university of Quebec (Formation à distance), and University of Montreal (Technologie éducationnelle), University of British Columbia (Technology Studies Education), and Royal Roads

(Distributed Learning). Degree programs in Communication and Media Studies were omitted due to the fact that these were more general in scope and not directly linked to Education. Similarly, degree programs in Educational Computing were omitted, since these have a longstanding history of their own that is connected to a different academic tradition. Finally, the analysis did not include undergraduate degree programs in Educational Technology due to the fact that none exist in Canada.

The selection of participants to interview was determined by representational and informational concerns. That is to say, efforts were made to interview educational technology faculty who were involved in the creation and discontinuation of programs at all the institutions being investigated. Participants were, therefore, educational technology faculty members selected according to the following criteria: (1) faculty member affiliation (present or past) to one of the Canadian universities which offered or offer degree programs in educational technology, (2) faculty member affiliation when the target academic programs were created, and (3) faculty member affiliation for the current year or the last year in which new enrolments were allowed.

The most senior current faculty members (or senior faculty members during the last year new students were admitted before a program was suspended) were interviewed within an educational technology degree program at each targeted university. Failing this, the second senior faculty member at the program's creation was interviewed. Failing this, the third senior faculty member at the program's creation was interviewed. In the event that there were two or more faculty members with senior rank, the faculty member with the most years of work experience at the targeted university was interviewed. The researcher also interviewed the senior faculty member at the point in time when the educational technology degree program(s) was created at each targeted university. This group of educational technology faculty form an occupational subculture that transcends organizational boundaries, developing their own cultural forms within the organization that are not public knowledge (Trice & Beyer, 1993).

## **Data Collected**

The primary researcher collected all data to ensure it was recorded in a consistent manner and with the aim of identifying themes and categories to be comparatively analyzed in order to "to generate or discover a theory" (Creswell, 1998, p. 56). Interview data was collected from educational technology faculty members who contributed to creating and discontinuing degree programs. Documents including academic program documentation, faculty participant curriculum vitae, and research on educational technology programs and professional associations in Canada was collected in order to contextualize the academic field of educational technology in Canada. Government documentation was also consulted.

*Interview Data.* Because the focus of the study was the development of academic educational technology in Canada, the first source of data for this study was 'insider' information on the academic field from individuals within the field in Canada. A minimum number of two interviewees for each of the 12 institutions were selected to ensure adequate coverage of each institution. In accordance with grounded theory methodology, interviews proceeded until no new information appeared to be forthcoming. One additional interview was added to the study to ensure adequate data saturation (Straus & Corbin, 1998). Consequently, 25 participants were interviewed to ensure adequate coverage of the academic field of educational technology in Canada. The interviews were semi-structured, primarily conducted by telephone, and recorded on audio tape.

From faculty member participants, the researcher gathered data on their backgrounds and viewpoints on how the academic field of educational technology evolved, their involvement with university degree programs at Canadian universities, summary of accomplishments (both their own and those perceived by others), and views on the future of the academic field by means of the following questions:

- Please tell me how you got involved in the academic field of Educational Technology.
- In your opinion, how did the educational technology program(s) evolve where you worked? What was your role?
- What do you consider to be the main contributions to the academic field of Educational Technology from you and colleagues in your program? (I.e., research accomplishments, innovations).
- What service projects have you and your colleagues been involved with within your university community?
- In your opinion, did external factors influence the academic field of Educational Technology? Which ones? (I.e., professional associations, government, private sector).



- Name five key events in the development of Educational Technology in Canada?
- In your opinion, what constitutes the academic community of Educational Technology? (I.e., which people, what fields).
- Describe how you feel others in the university perceive the academic field of Educational Technology? How do others in the community you serve perceive it?
- What professional associations are you affiliated with? What associations would you rank as the most important to the development of Educational Technology in Canada?
- What are your thoughts about the future of the academic field of Educational Technology in Canada?

*Document Data.* In addition to insider information, another source of information about the development of the academic field of Educational Technology was documentation that educational technologists have produced. These documents explain how educational technologists view themselves and their academic fields. Sources about an academic field typically include their records of research, such as refereed journals and documentation about research. Sources also include the documented history of academic programs and professional organizations, such as newsletters, program guides, and bulletins (Trice & Beyer, 1994). To capture information about participants' contribution to the academic field, the researcher reviewed curriculum vitae (provided by participants prior to interview) in relation to interview data from the 25 participants. To capture information about the history of academic degree programs, program information was examined from 12 universities in Canada, along with existing research on academic programs.

Program information collected included type of degree, degree name, courses offered, required courses, faculty contingency, and number of other programs offered within the academic department or unit. Where possible, program information was sampled at five year intervals beginning with 1969-1970 to 2004-2005 in addition to sampling program information from 2006 and 2007. Public access to program information from previous years was available during normal business hours at the university archives on site or through telephone request. The review of educational technology program information consisted of documents derived from the 12 institutions studied. Document data collected about the academic field also included research on educational technology programs and professional associations in Canada. Finally, provincial and federal government documentation was also consulted in relation to the interview data from participants. This included annual reports and documentation from government agencies connected to educational technology projects. In total, approximately 200 documents were selected for review to ensure document manageability. All documents utilized in the study were catalogued.

## **Data Analysis**

The study was guided by grounded theory methodology for analyzing the data where themes and categories were developed in phases based on Strauss and Corbin (1998). Overall, data analysis strategies supported the study by focusing on self-report interview and documentation data revolving around the development of the academic field of educational technology in Canada. First, initial coding was performed on interview data and documents. In initial coding, concepts were compared with one another and those that seemed to pertain to a similar phenomenon were grouped into categories (open coding). An analysis was performed on interview data and documents from each institution separately using open coding, which entailed the forming of "initial categories of information about the phenomenon being studied by segmenting information" (Strauss & Corbin, 1998, p. 57). The second tier of coding (axial coding) created connections between main categories and sub-categories by scrutinizing individual categories in terms of their underlying conditions, context, interactions, and consequences (Strauss & Corbin, 1998). An analysis was performed on interview data and documents from the 12 universities together to identify the conditions that influence the phenomenon and also the consequences of it. Finally, a model concerning how the phenomenon is perceived by the participants was developed (Strauss & Corbin, 1998). The result was a grounded perspective intended to explore current challenges and future developments the academic field of Educational Technology in Canada reached its current state.

## **Trustworthiness**

This study was intended to generate new knowledge of the field, while ensuring that the researcher entered the study in as bias-free a manner as possible (Denzin & Lincoln, 2000). To make sure that the conclusions of this study could

be relied upon by other researchers, I employed a number of procedures to ensure the trustworthiness of the study. An initial draft of interview questions was reviewed by educational technology faculty members from the University of Laval and Concordia University. Subsequently, modifications were made to the interview questions to ensure the effectiveness of the instrument and the usefulness of the information gathered. Second, to ensure the accuracy of evidence, data was triangulated. Multiple sources were used, including several interviews and documents, so that descriptions of phenomena were not merely the opinion of a single participant or a group of participants, but also validated by the documents (Denzin & Lincoln, 2000). Third, member checks were performed throughout data collection and analysis as a second procedure to ensure the accuracy of evidence, meaning that all interviews were returned to study participants for verification and modification if necessary. In addition, institutional profiles were also subject to member checks through verification of profile content with the senior faculty participant from each of the 12 institutions included in the study. Finally, to address the possibility of researcher bias, the researcher participated in an ongoing exercise of memo writing throughout the study. In these memos, reflections were recorded about the study, thoughts about the interviews and phenomena observed, and opinions stated by participants. This allowed the researcher to discern possible problems and biases that may influence the interpretation of findings.

## Findings

A qualitative study was carried out to gauge the state of the field, current issues, and possible future developments. The main category discovered in all twenty-five documents (100%) revealed that all participants perceived there to be challenges connected to the academic field of Educational Technology in Canada. This main category contains seven key sub-categories, namely, identity, standardization and professionalization, university politics, external influence, competition, funding, and teaching and learning. Overall findings are represented below in Figure 1 and discussed.



Figure 1: Main Category and Sub-Categories

### Identity

Identity in the academic field of Educational Technology was the most predominant category emerging from interview data and review of the professional literature analysis. An examination of a total of sixty-five text units from sixteen out of twenty-five documents (64%) revealed distinct challenges revolving around the conceptualization of the field and its parts, along with public recognition and reputation building.

One way to find out about a field is to look at how its members conceptualize it and agree on shared meaning(s). In carrying out this study, there appeared to be no unified conceptualization of the academic field within the

professional literature and faculty viewpoint. One popular viewpoint was that Educational Technology was a field about engineering effective learning systems with the accent on the learning (24). Another viewpoint was that Educational Technology was not a field and that it was better served as a set of perspectives. A third viewpoint was that the field of Educational Technology and its academic programs should be interdisciplinary programs with people coming from contributing disciplines and that the field should be defined in broad terms:

There is an increasing critical mass of scholars and students who see educational technology in a larger context and who want to do more critical things with educational technology and look at things in and around things like cyborgs, cyberculture, and intellectual property issues. (20).

A number of experts speculated on why defining the field was problematic. One expert indicated that it could be the lack of a subject matter focus that makes it difficult to define(1). Another expert pointed out the problem of being too broad. He stated, “If you define technology as a process, then Educational Technology includes all dimensions of education” (6). A third expert believed that the concept of Educational Technology needed to be ‘demystified’(7). Another expert believed that there is a problem in terms of the language that is being used and the way people feel about the areas they have ownership over (13).

In considering key areas within Distance Education and Educational Technology, important relationships were highlighted between the two. When probed about the connection between the academic areas of Distance Education and Educational Technology in Canada, there were mixed perspectives. In one camp there were those who viewed Distance Education as a sub-set of Educational Technology. Under this view, Distance Education involves a distinct set of access and adult education concerns unique from educational technology applications in campus settings (1). Some experts considered Distance Education to be the most popular area of application of Educational Technology in education but that Distance Education specialists do not typically realize that they are a particular breed of educational technologist. A second camp viewed Distance Education and Educational Technology as separate fields. Most experts indicated that the two fields represent two solitudes which have basically ignored each other and that their relationship has not been necessarily synergistic nor even connected in past and current developments.

There are also a number of field identity challenges concerning reputation building and gaining public recognition. The growth of educational media in the 1960s and 1970s was one area connected to successful reputation building. One expert reflected, “Media people like myself became important because we could make things run. We were willing to travel from person to person and straighten out problems with their media equipment” (2). Gradually, the field moved from instructing media to actually influencing the teaching-learning process.

Starting around 1980, the introduction of instructional development helped to bring audiovisual media into an academic configuration. This gave educational technologists visibility and a wedge into mainstream Education with greater respect within the academic community. In the 1990s, there was additional recognition of the field through greater acceptance of e-learning as a legitimate form of learning.

From 2000 to 2006, additional recognition in the field a growing popularization of technology integration in education led to many people putting their course materials online. Beyond reputation building activities in selected areas, there appeared to be an ongoing struggle for public recognition. This is partly due to the broad focus of the field promoting scholarship and teaching across disciplines, along with an emphasis on technology and innovation. One expert remarked that the community is very eclectic but there is not a lot of respect for the body of knowledge accumulated in the field by people too focused on technology and innovation, “Because technology continuously offers new toys to play with and because of vendor hype, is exciting. It seems to lend some prestige to those associated with the latest techno excitement” (24).

## **Funding**

Funding issues led to a number of challenges discussed by experts in a total number of 78 text units in 16 out of 25 documents (64%). The main challenges identified were linked to selective funding areas, university funding complexities, and funding cutbacks. Part of the funding challenge in the academic field of Educational Technology arises due to funding availability in selected areas. One expert remarked how faculty members have had funding from different sources to promote IT and education integration projects (9). Another expert commented that there are

often targeted funding projects where money was made available to carry out a specific objective such as the wiring of schools (11).

There is often money available for doing research on technology but funding sources are often less willing to put money into research on pedagogy associated with the technology (13). Experts were quick to point out that there is more money in the scientific fields than in basic fields such as education (23).

Funding problems were also present at universities with department based research centres generating significant income for the university which supported the research infrastructure. One expert indicated that there was a constant tension because research centers want some of the funds back to cover costs but that universities often want to give less (21). Another remarked, "We just had a meeting with the dean about the fact that we are generating tons of revenue but we are not hiring the full time faculty that we need (20)."

Funding challenges appeared to influence academic units as indicated by one of the experts. Educational Technology programs and departments typically did not have control over the resources that came to our department and have to negotiate funding with department chairs and faculty deans managing dozens of other programs (21). Funding cutbacks emerged as an important challenge discussed by experts. For instance, one expert pointed out how his department "flew apart because of frozen budgets" (3). Loss of funding also appeared to influence professional association activity. Ongoing drop-offs in funding and membership in AMTEC and CADE helped to motivate the two associations into looking at amalgamation (3). There are additional challenges with funding agencies such as funding irregularity and conflicting stakeholder goals.

Some experts believed that funding agencies like Industry Canada are only interested to innovation. One expert remarked, "In the past the government sponsored projects by throwing money at them keeping their fingers crossed that they will reap results. In some instances a pilot was set up and money continued to come when there no reason and other times a project that was working suddenly got its funding stopped" (7). This had drastic effects on the sustainability of many projects. One expert who received funding agency support indicated, "When the funding ran out the center petered away" (7).

A second theme addressed conflicting goals and limited vision. One expert stated that, "you have other groups in the political sphere like Industry Canada and different government groups that have promoted the use of educational technology in certain ways and they have all had different goals" (7). This creates a challenge for researchers to figure out what are the key funding agency goals. One expert indicated, "There was always the federal and provincial government money attached to figuring out whether the newest innovation worked" (18).

## **University Politics**

The second most predominant set of challenges identified by experts focused on university politics. An examination of a total of sixty text units from fifteen out of twenty-five documents (60%) revealed distinct challenges revolving around the acquisition of limited resources, faculty hiring, program organization, and public perceptions of Educational Technology within universities.

In university settings there is often a 'push back' when people want to legitimize the study of learning because it is not discipline based (7). University influences affected the creation and modification of programs. As indicated by one expert, "The only hold up is that they [University administration] do not feel we have the human resources and funding to handle another graduate program in conjunction with what we are doing now in the master's area"(3). In addressing issues of program modification and institutionalization, experts also noted that changing program names was often a challenge because once a program is named there is an affinity among the graduates for the name, which makes it even more difficult to change it (8).

University control over faculty hiring was another issue linked by experts to university politics. One expert stated, "We sometimes hired people for the wrong reasons and I fought tooth and nail against this" (24). Experts from the majority of universities studied commented on the lack of tenured positions and a saturation of sessional part-time faculty members to teach course sections with few privileges and little stability. In many cases, program development was held up due to the nature of faculty hiring. This was particularly problematic at institutions where

there were not enough qualified tenured faculty to supervise students in doctoral programs. At some institutions, freezes on faculty hiring were common (2). University organization of educational technology programs within education departments was considered to be another challenge. As stated by one expert, "Educational Technology has restricted its vision of the domain by maintaining programs in university departments of education." (11).

Regarding public perceptions within the university, some experts believed they were often perceived as a utility to others within the university. In order to do more with less, a lot of academics felt that educational technology would alleviate some of the budget freezes that had been occurring (3). Many experts indicated that faculty outside of Educational Technology viewed technology as a tool to help them teach large classes and a way to make their lives more convenient in terms of records management, testing, etc. (2). Attitudes ranged from those who saw it as a new tool with a lot of promise to those who viewed it as the replacement for pencil and paper (19). Under this view educational technologists are perceived as technicians and academics look upon educational technology academics, not as academics, but as "techies"(1). One expert commented, "Many people, even within our own faculty, believe that educational technology is only about computer technology. Thirty years ago, they would have said it was about slide and movie projectors" (11).

### **External Influence**

External influence led to a number of challenges discussed by experts in a total number of 68 text units in 14 out of 25 documents (56%). A number of challenges were identified regarding influences from the provincial and federal government, the private sector, and the United States. The provincial government created challenges connected program development. There are political influences through provincial regulations preventing universities to compete with other universities by offering similar programs (1). The private sector appeared to have an influence, especially companies with interest in educational technology, training, and publishing. This includes Canadian National, Canadian Pacific, Air Canada, Bell Canada, banks, hardware and software makers, and publishing firms (2). This created challenges concerning private sector control over decision-making within the academic field. For instance, companies like WebCT can offer course management systems at a low introductory rate and increase charges dramatically once implemented (23).

There are also commercial interests from those who want to sell hardware and software to schools and to demonstrate that it is good for their financial interest (11). Some experts noted challenges of financial dependency when private sector interest disappears as was the case in Quebec in the early 1990s when private sector influence dropped off due to takeovers, internationalizations, and companies spinning off their training departments (2).

American influence represented challenges for some experts in reference to the Canadian development of the academic field of Educational Technology. American influence was discussed in terms of government, programs, faculty, and the US academic field. One expert linked the American influence to technology developments in the US that rippled over into Canada along with educational technology faculty graduating from American universities. Experts offered a number of reasons for the strong American influence including Canada's proximity to the US, and the critical mass of American economics. Given the strong American influence on the academic field, a number of experts were concerned with the Canadian contribution to the academic field of Educational Technology. One expert remarked, "There really was a distinctive Canadian contribution in the 1960s and 1970s that died away" (2). Another expert commented, "We as Canadians have trouble thinking like a country instead of the sum of provinces. We are caught in a provincial mode and have no pan-Canadian initiatives" (20).

### **Standardization and Professionalization**

Standards and professionalization represented another area of concern within the professional literature and in fifty-one text units from ten out of twenty-five documents (44%). A number of experts speculated on why creating standards was problematic. One expert indicated that because Educational technology was continually shifting and changing, standards were difficult to define" (1). Rempel, Montgomerie, and Szabo (1998) noted how standards for Instructional Technology (IT) practitioner competency requirements were evolving and changing. This ongoing challenge concerned the professionalization of Educational Technology and was particularly salient in Quebec.

In Quebec, developments from professional associations began in the 1960s when the audiovisual focus turned to Educational Technology focus. By the 1980s, ADATE began appealing to the Quebec government to be recognized as a profession. The “Manuel Professionnel de l'ADATE was circulated in more than 2000 copies” (6). Following the publication of ADATE’s professional manual, the Quebec Office of Professionals refused a request to recognize Educational Technology as a profession because its specialists were within the teaching profession, which was not a recognized professional order”(6). Proponents of field professionalization struggled for greater autonomy and a say over the basic training of those who want to enter the profession, like engineers or doctors (6). Some experts believed that there is a pressing need for someone to wave the flag of Educational Technology as a specialty in order to overcome limiting factors such as identity branding challenges and competition.

## **Competition**

To a somewhat lesser extent, various forms of competition creates challenges within the academic field. Experts generated a total number of 21 text units in 10 out of 25 documents (40%) on competition. The main challenges revolved around university competition, program competition, and tension between professional associations. Within universities there was competition within academic units for limited resources and also for content, particularly when Educational Technology is applied to areas like primary or secondary schools, tertiary education or the workplace. As indicated by one expert, people from education view this as an encroachment on their work in primary and secondary schools and on Adult Education” (24). This was a concern because there is also the ‘gobbling up’ of one academic area by another (2). At the same time, there are a lot more groups coming into play, who some people would say are intruding on our territory (22). There was notable competition over the limited number of faculty positions at some universities by groups with more political power.

Competition between universities was a challenge as it related to program creation and development. For instance, some provincial governments have strict policies concerning the replication of existing programs and competition between universities, creating barriers to new program creation. There were also problems with competition between already existing programs, particularly in the 1990s when program clientele dropped off because some universities began offering shorter masters’ degrees (2). On a broader scale, conflicting perspectives from competing professional associations was believed to be a challenge to developing a strong Canadian professional associations. One of the big tensions within AMTEC concentrated on whether they would lose the K to 12 practitioner and media people from the association after the merger with CADE took place (22).

## **Teaching and Learning**

To a somewhat lesser extent, a number of teaching and learning challenges within the academic field generated total number of 20 text units in 10 out of 25 documents (40%). The main challenges concerned faculty interest and expertise, along with hiring practices. Regarding faculty interests, there was notable concern over the presence of a ‘tool fetish’ approach to Educational Technology and a fair percentage of academics and students of educational technology with a narrow view of educational technology as nothing more than the sum of ICT (20)”. The issue as described by one expert was, “The fact that technology was popular and that people developed the impression of having expertise on the subject reduced the meaning of Educational Technology to a domain that deals with instructional tools “(25). One expert remarked how ‘educational technophiles’ do not understand that they will never be computer scientists, engineers, or television producers (24). He noted , “In as much as our field is one of engineering systems to help people learn and to apply learning in ways that are valued, many of the faculty to this day cannot get past machinery, baubles, and pictures” (24).

Another major concern revolved around faculty hiring and teaching methods. One expert stated, “Some of the old-fashioned methods may come from people who have been teaching for twenty years and have never evolved beyond a correspondence course style or telephone based methods of teaching” (1). A second expert contributed a similar perspective, “I think that people who are in many of the current academic positions are not contributing to our greater glory to be quite honest (24)”. He elaborated, “The majority of the faculty had come from older traditions and were essentially just teacher trainers papered over” (24).

The level and breadth of expertise was also perceived to be an issue. As indicated by one expert, “We cannot simultaneously be good instructional designers, computer based design experts, organizational design experts, management experts, and incentive system experts” (21). He continued, “We suffer from the challenge of being a jack of all trades and a master of none because there are so many things associated with addressing learning and performance needs that there is no way to be an experts in all areas” (21).

## The Future of Educational Technology

Regarding the future of the academic field of Educational Technology, Beckwith (1988) recommended a stronger emphasis on transforming learners and learning processes through a systemic approach. Beckwith's (1988), “The Future of Educational Technology” was a significant article in the field positing three traps the developing field should avoid: Compromised integrity (defining learning and instructional problems in terms of the latest innovation rather than focusing on educational goals), adherence to status quo (emulating established professions to gain credibility rather than focusing on recommending significant changes), and solidification (investing intellectual energy in short-term survival and routine applications rather than focusing intellectual processes on systemic approaches to advance long term goals to benefit society). Other contributions to the professional literature in this area were provided by Mitchell (1989), “The future of Educational Technology is past” and follow-up responses published in the *Canadian Journal of Educational Technology* from Andrew Agostino, Don Beckwith, Ivor Davies, Philippe Duchastel Michael Hannafin, Stephen Kerr, Gerald Torkelson, and William Winn (Canadian Journal of Educational Technology, 1989).

When asked to provide views about the possible future development of the academic field of Educational Technology in Canada, key area discussed included academic programs, professional organizations, and movements and trends. Table 1 illustrates experts' hopes and fears on key areas of future development.

*Table 1: Experts' hopes and fears on future development in Educational Technology*

<b>Future Development</b>	<b>Hopes</b>	<b>Fears</b>
Academic Programs	New doctoral programs New certificate programs	Faculty shortage, university politics and funding cutbacks, and limited resources prevents new program development
Professional Organizations	Professional association partnerships create greater cohesion within field	Professional association partnerships do not affect cohesion within field
Movements and Trends	Strategic government-university and private sector-university partnering advances field in expensive areas involving new technology	Strategic government-university and private sector-university partnering causes field to retreat.
	Successful renewal of tenured faculty.	Increase in part-time faculty hiring without renewal of tenured faculty.

Regarding the future of programs, many experts mentioned hopes for expansion. It is expected that a proposal for a doctoral degree in Distance Education at Athabasca University is going to go through in 2008, making it the first doctoral program in Distance Education delivered at a distance (1). The University of Saskatchewan has a doctoral program proposal submitted to the College of Graduate Studies (3). Other universities like, Laval University, are developing a certificate programs. Future plans were not without perceived barriers to overcome. At the University of Calgary, there are plans for expansion but there are many issues, such as availability of faculty (12), economic pressures forcing universities to make cutbacks (2), and continued competition from other fields (22). For instance, information science groups, human-computer interface groups, and a growing number of informatics programs are interested in many areas that overlap with instructional design and technology. One expert mentioned, “There are a lot more groups coming into play, who some people would say, are intruding on our territory (22).

The future of professional organizations and associations was also a key theme. In reference to the professional associations, a number of comments were made. One expert indicated, “These fields have diverged in the past but the joining of CADE with AMTEC will bring it all back together” (3). The merger is expected to provide a more cohesive venue for ideas by bringing these people together (12). If and it works, it would be the first organization in

Canada to pull key areas of the field together and provide an opportunity for growth” (22). Some experts, however, believed that a professional association merger would make no difference (11). It was also noted that attempts to merger the professional associations were not successful in the past (3).

Future movements and trends were also discussed by a number of experts. Experts speculated on the growing trend in private sector participation. One expert stated, “I see a lot of the current computer labs sitting empty and not being used for educational purposes. I see this void being filled by hardware and software makers and by publishing firms which I think will reinvent themselves just as libraries are reinventing themselves” (7). Some experts believed the future growth of Distance Education and continued expansion of online learning would bring convergence of video and online education (1) and more universities with technology service centers (21). Others speculated that the past would be relived with no major developments emerging much like reinventing the wheel from generation to generation with new educational media and the same types of old research done with each medium (1).

Future scholar renewal was also discussed by a number of experts. According to one expert “There is a crossroad that we are at and it is extremely important for future scholars in the job market (20). Experts suggested a hopeful future for the field provided that new faculty produce quality research (16) and valorize research on the history and development of Educational Technology (25). Concerns over faculty hiring freezes and a reduction in tenured faculty hiring were expressed by experts. Partnering initiatives were also associated with future developments, especially inter-university, university-government, and university-private sector partnering (6).

### **Strategic Planning for Educational Technology in Canada**

Regarding the connection between experts’ views and the professional literature, findings indicated a high degree of overlap in key challenges identified. Knowledge of the abovementioned challenges (identity branding, standardization and professionalization, external influence, funding, competition, teaching and learning) can be applied to current strategic planning to leverage educational technology academic programs as a means of advancing evolving the field in Canada.

Given the complexity of the field (tool level, systematic level, systemic level) and the complex set of challenges facing it, there is warrant for strategic planning addressing the organizational complexity of Educational Technology program planning. It is therefore, necessary, in the application of strategic planning to address identified challenges and to contextualize program planning within the university environment (internal environment) embedded within a larger social and political context (external environment). In line with Cook (1990) and Kaufman (1991), a grounded strategic planning model emphasizes the context from which challenges were discerned and the future importance of decisions made in the present. See Figure 2.

In Figure 2, the strategic planning process begins with a preliminary intra- and inter-institutional inquiry (context), the creation of planning committee, and a systemic diagnostic of the internal and external environment. In a complex applied field like Educational Technology, a preliminary inquiry work can take months. In the present study, data collection and analysis of relevant professional literature and insider expertise allowed for the identification of key challenges emerging from the internal and external environment. From this, strengths, weaknesses, potential threats and opportunities can be identified in order to discern shared needs and interests from all stakeholders involved. This strategic planning approach accommodates interests and needs from different stakeholders. Processes are designed for achieving consensus and thus an effective decision making in line with recommendations from Bryson (1990). This is a key concern in Educational Technology as a field with competing conceptions of its meaning and significance. As indicated by one expert, “Educational Technology makes a lot of sense but the whole field of Education is so fragmented that it is hard for one group to see the benefits of another group” (14)”.

Next, strategic partnering, resource allocation, and the identification of objectives and strategies are listed. These steps are important in ensuring that the necessary conditions for supporting Educational Technology initiatives are available. According to one expert, “The only way to have a distinct Educational Technology is if you have a lot of money and good partnerships with industry to push” (14). This is partly due to the high cost of research in Educational Technology and the high need for resources and support people (18).



Finally, the implementation and evaluation of strategic initiative(s) are iterative processes followed by a return to further data collection and analysis, and changes to planning committee if needed. This is particularly important in this present context where there is no clear and unified conception of Educational Technology and a low public awareness the field. As indicated by one expert, “We have never been able to give anybody very definitive answers about things. There is no general body of knowledge that all people can relate to” (18).

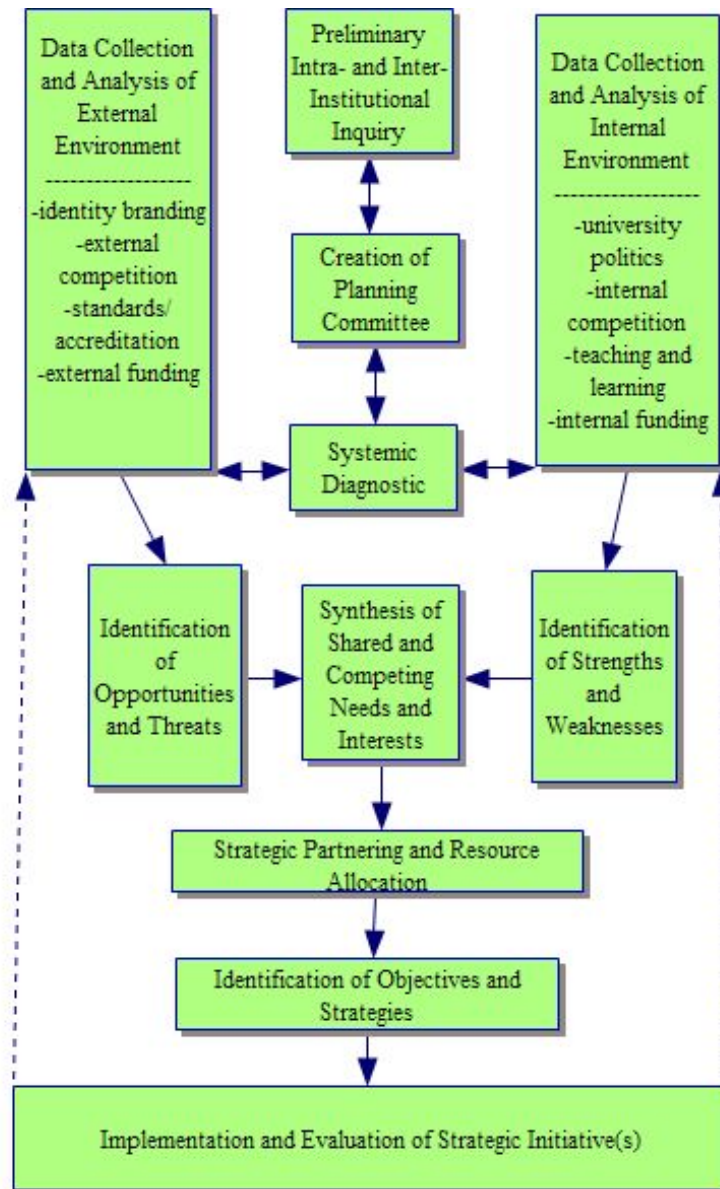


Figure 2: Grounded Strategic Planning and Evaluation Model

Moreover, during the evolution of a complex applied field, stakeholders and stakeholder interests can change. For instance, faculty retirement trends in Educational Technology during the 1990s led to substantial program restructuring and, in some cases, program closures within Canadian universities (Luppicini, 2006). It is therefore, necessary to be able to engage in ongoing inquiry and analysis. In line with Hlynka and Nelson (1985), the abovementioned strategic planning processes are multi-level and iterative, allowing for a high level of synergy between established approaches to educational technology (the tools approach, the systematic approach and the systemic approach) in addressing current challenges and possible future development within the field.

## Discussion

Educational Technology is currently at a crossroads in its evolution due to a number of important challenges discussed in this article. This article presented the collective thoughts and integrate documents concerning challenges and possible future development within academic field of Educational Technology in Canada. The first part of the section analyzed the main category (challenges) with seven sub-categories, which emerged during the analysis: identity branding, standardization and professionalization, university politics, external influence, competition, funding, and teaching and learning. The second part of this section, highlighted experts' views on possible future developments within the academic field of Educational Technology connected to academic programs, professional organizations, and movements and trends. Based on a synthesis of expert interview data and relevant professional literature, strategic planning processes emerged delineating major challenges and possible future developments.

Two recommendations are offered to scholars and practitioners in Educational Technology. Following from study findings and speculative strategic planning processes, the first recommendation is aimed at leveraging the identity of Educational Technology in Canada. It is strongly suggested that greater attention be invested in identity branding activities within the field. This implies that a greater emphasis be placed on retrospective studies of the field and best practices. This may help lend greater respectability to the field if historical developments within the field of Educational Technology become valorized. At the same time, this will assist Educational Technology scholars and practitioners avoid current challenges of reinventing the wheel by promoting a greater self-awareness in the field of past events and current developments to help guide future activities (22).

The second recommendation is aimed at leveraging Educational Technology in Canada through greater attention to partnering possibilities, their threats and potential opportunities (see Figure 2). Leveraging the public image and market appeal of Educational Technology is paramount to the survival of the field. According to Rempel, Montgomerie, and Szabo (1998), "The future survival of IT programs may well depend on their ability to develop both instructionally sound and commercially viable programs, including programs that could be targeted to specific short term market needs." Similar recommendations on the importance of public marketing within academia are offered elsewhere by Cook (1990), Kaufman and Herman (1991), and Kottler and Fox (1985). This connects to key challenges discussed in this study. For instance, in the case of standards and professionalization, experts perceived future benefits in collaborating to establish a pan-Canadian standard(20) and creating standardized definitions and cultivating a national dialogue on about what experts are doing in the field(18). The potential of partnering to leverage Educational Technology opportunities in Canada has be balanced with potential partnering threats as well from partnering with industry, military or universities. Gagne (1987) remarked that instructional designer knowledge and expertise must be guarded from contamination, and not be compromised by the various influences of the marketplace.

Beyond the obvious implications for continued research on the field of Educational Technology in Canada, the grounded model posited generated in this study may provide useful insight for more practical forms of program evaluation suitable for applied fields of study. While traditional work in curriculum theory and evaluation focuses on deciding what should be taught, why, and to whom (Jackson, 1992; Pinar, Reynolds, Slattery, & Taubman, 1995), this study suggests that applied fields (like Educational Technology) require a special understanding of complex relations affecting the fit between program planning and management, market demands, and student interests within the context of society. The grounded perspective generated in this study may be tested on other applied university programs such as Science and Technology Studies, Applied Human Sciences, and Information Sciences.

## References

- Beckwith, D. (1988). The future of educational technology. *Canadian Journal of Educational Communication*, 17(1), 3-20.
- Bernard, R., & Lundgren, C. (1991). Educational technology in transition: A study of the years 1968-1989. *Canadian Journal of Educational Communication*, 20 (2), 125-140.
- AMTEC (2008). *Canadian Network for Innovation in Education*, Retrieved, July 15, 2008 from <http://www.amtec.ca/index2.html>.

- Bryson, J. M. (1990). *Strategic planning for public and nonprofit organizations*, San Francisco, CA: Jossey-Bass.
- Canadian Journal of Educational Communication (1989). Special issue. *Canadian Journal of Educational Communication*, 18 (2).
- Clark, R.E. (1984). "Where's the beef?": A reply to Heinich. *Educational Communications and Technology Journal*, 32(4), 229-232.
- Cook, B. (1990). *Strategic planning for America's schools*, Arlington, VA: American Association of School Administrators.
- Creswell, W. J. (1998/2006). *Qualitative Inquiry and Research design: Choosing Among Five Traditions*, Thousand Oaks, CA: Sage.
- Davies, I. (1984). Instructional development: Themata, archetypes, paradigms and models. In R.K. Bass & C.R. Dills (Eds.), *Instructional development: The state of the art, II*, Dubuque, IA: Kendall/Hunt, 9-17.
- Davies, I. (1978). Prologue: Educational technology: Archetypes, paradigms, and models. In J. Hartley, & I. Davies (Eds.), *Contributions to an Educational Technology*. London: Butterworths.
- Denzin, N., & Lincoln, Y. (2000). *Handbook of qualitative research* (2<sup>nd</sup> Ed.), London: Sage.
- Finn, J.D. (1964). The Franks had the right idea. *NEA Journal*, 53(4), 24-27.
- Heinich, R. (1984). The proper study of educational technology. *Educational Communications and Technology Journal*, 32 (2), 61-87.
- Hlynka, D. (1994). Marginalizing significant others: The Canadian contribution to educational technology. *Educational Technology*, 34 (2) 49-51.
- Hlynka, D., & Nelson, B. (1985). Educational technology as metaphor. *Programmed Learning and Educational Technology Journal*, 22 (1), 1-15.
- Jackson, P. (1992). Conceptions of Curriculum and Curriculum Specialists. In Philip W. Jackson (Ed.), *Handbook of Research on Curriculum: A Project of the American Educational Research Association*, NY: Macmillan, 3-40.
- Kaufman, R., & Herman, J. (1991). *Strategic planning in education: Rethinking, restructuring, revitalizing*, Lancaster, PA: Technomic.
- Kottler, P., & Fox, K. (1985). *Strategic marketing for educational institutions*, Englewood Cliffs, NJ: Prentice Hall.
- Luppici, R. (2006). Connecting with the development of Educational Technology in Canada. *Paper Presented at the Association for Educational Communication and Technology (AECT) Conference*, Dallas, USA.
- Luppici, R. (2005). A Systems Definition of educational technology in Society. *Educational Technology and Society*, 8 (3), 103-109.
- Miles, M., & Huberman, A. (1994). *Qualitative data analysis: A sourcebook of new methods* (2<sup>nd</sup> Ed.), Thousand Oaks, CA: Sage.
- Mitchell, P. (1975). The discernible educational technologist. *Programmed Learning and Educational Technology*, 12 (5), 306-325.
- Mitchell, D. (1989). The future of educational technology is past. *Canadian Journal of Educational Communication*, 18 (1), 3-27.
- Pinar, W., Reynolds, W., Slattery, P., & Taubman, P. (1995). *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*, New York: Peter Lang.
- Plomp, T., & Pals, N. (1989). Continental European perspectives. In M. Eraut (Ed.), *The international encyclopedia of educational technology*, Brighton: Pergamon, 51-54.
- Reiser, R. (1987). Instructional technology: A history. In R Gagne (Ed.), *Instructional technology: Foundations*, New Jersey: Lawrence Erlbaum, 1-34.
- Rempel, P., Montgomerie, T.C., & Szabo, M. (1998). IT Roles and Competencies: Are We Prepared for the 21st Century. In Proceedings of ED-Media/ED-Telecom 98, Charlottesville: AACE, 1188-2193.

- Silber, K.H. (1970). What field are we in, anyhow? *Audio-Visual Instruction*, 15(5), 21-24.
- Silber, K.H. (1978). Problems and needed directions in the profession of educational technology. *Educational Communications and Technology Journal*, 26(2), 174-185.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques*, Newbury Park, CA: Sage.
- Torkelson, G.M. (1987). Perspective: Theoretical bases for research in media. *Canadian Journal of Educational Communication*, 16 (1), 3-9.
- Trice, H., & Beyer, J. (1993). *The cultures of work organizations*, New Jersey: Prentice-Hall.
- Yeaman, A., Hlynka, D., Anderson, J., Damarin, S., & Muffoletto, R. (1996). Postmodern and poststructuralist theory. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology*, New York: Simon and Shuster Macmillan, 226-252.

## Appendix I: Perceived Challenges in the Evolving Educational Technology Field (N=25)

Current Challenges in the Evolving Educational Technology Field	
• Large number of professional associations with overlapping interests	Faculty in the field coming from older traditions.
• Faculty not necessarily skilled or even interested in technology	Individuals in the field contributing very little to academic institutions.
• Old-fashioned methods from people who have been teaching for twenty years	Individuals in academic positions not contributing to the field.
• Teaching program competencies for dealing with market expectations.	University struggles attracting new faculty members.
• Migration of faculty to the graduate courses and contract people teaching B.Ed. courses	Poor hiring practices.
• Educational technology programs becoming digital diploma mills.	Technology usage in Canada lagging behind other countries
• Course design by full -time faculty with part-time faculty teaching	Part-time collective agreement blocking part-time program appointments of international scholars.
• Too many graduate students in a graduate course	Inconsistencies in government funding for pilot projects.
• Increase in part-time faculty numbers with no privileges or stability.	Lack of funding and support for suitable equipment.
• Lack of systems thinking to address real-world problems within societal systems.	University debates concerning affiliate research center support.
• Rising need of non-instructional approaches and solutions	Continual effort to reinvent the wheel without the realization that much of the work was already done ten or twenty years ago.
• Increased demands to work at the organizational and team level	Individuals in charge of technology not aligning themselves with public interest.
• Challenge of being a jack of all trades and a master of none	Disconnect between the kinds of theory and research that is being practiced.
• Student attrition in large graduate programs with minimal intervention	Research is often not definitive and does not stand up to intense scrutiny.
• Students come from diverse backgrounds and have different demands	Reinforcement of stereotype that all educational technologists deal with are tools.
• Allowing students to slide through programs without critical evaluation	Low marketing budget for raising public awareness of field.
• Students without predisposition to theoretical foundations.	Persistent entrenchment of traditional instructional design models.
• Student resistance to some required courses	Overemphasis on technology as tool perspective from Educational Technology faculty.
• Research difficulties in gathering participation, getting consent, etc.	It is what has held Educational Technology back.
• Maintaining enough faculty and research activity to support a Masters and doctorate program	Educational Technology is not widely appreciated within universities.
• Hostility from colleagues outside Educational Technology	Acquiring enough resources and funding to create new programs and manage existing programs.
• They get stuck on the media and the technology.	Faculty hiring freezes
• Private sector hype about new technology benefits.	Difficulties bringing together people interested in the impact of technology across the university.
• Short-term prestige for individuals associated with the latest technological fad.	University opposition to Educational Technology efforts to legitimize the study of learning because it is not discipline based.
• Educational technologists attempting to emulate computer scientists, or engineers.	Problems with university resource sharing and competition.
• Pursuit of individual research and work that does not advance the field.	Inconsistencies in program on paper and what is implemented.
• Lack of convergence between Educational Technology and Distance Education	University hiring of low paid part-time faculty.
• Lack of clarity among educational technologists in describing their field	Tendency of Ed. Tech. Masters programs drawing away resources from PhD Programs.

***ISSN: 1436-4522 (online)***  
***ISSN: 1176-3647 (print)***

***<http://www.ifets.info>***